Opening a Japanese Restaurant in Toronto

Introduction

Background

Toronto is the capital city of the province of Ontario, one of the largest cities in Canada by population with 2,731,571 residents as of 2016. The majority of the Ontario residents live the the Greater Toronto Area (GTA) which makes it Canada's most populous city. With the wide diversity in Canada and in the City of Toronto, there are many restaurants that offer almost every cuisine that exists on the planet. However, considering the increasing popularity of the oriental food, it is always a best idea to open an authentic Japanese restaurant.

Problem

Opening a new restaurant in the city would require a good understanding of the geography, neighborhoods and diversity of the city and the distribution of the restaurants on each side of the city that ranges from east to west and the center.

Interest

Our study would require filtering all the venues to only work on venues that serve food, fast food or are just diner places or restaurants. Also the information will be filtered in order to see how many Japanese restaurants are in each area. Our study should decide which area would be best to open an Japanese restaurant.

Data

Data sources

For the Toronto neighborhood data, I will use the following Wikipedia page: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M. This has all the details I need to explore and cluster the neighborhoods in Toronto. I will scrape the Wikipedia page by using BeautifulSoup and wrangle the data, clean it, and then read it into a pandas dataframe. In this way it will be in a structured format.

The geographical coordinates of each postal code will be read in a pandas dataframe as well from: https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDeveloperSkillsNetwork-DS0701EN-SkillsNetwork/labs_v1/Geospatial_Coordinates.csv.

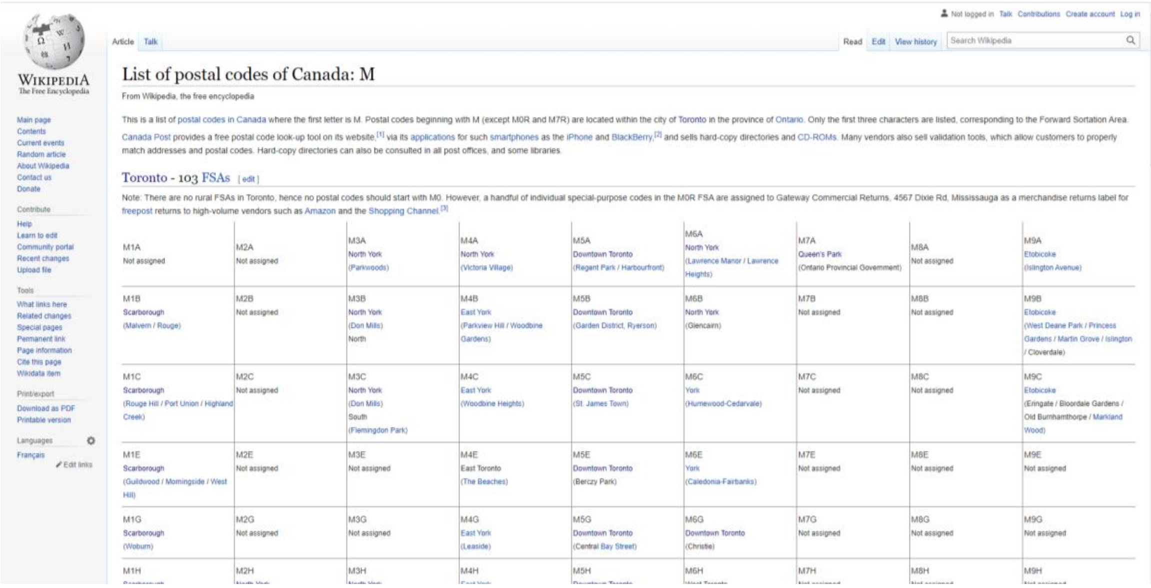
Data cleaning

The both dataframes obtained (the one with the neighborhoods and the one with the geographical coordinates) will be merged together so that each neighborhood has its geographical coordinates assigned.

I will use Foursquare APIs to fetch the data from all the venues in Toronto. Then I will filter the data to get only the restaurants/diners or food related venues to work on. The venue data obtained will help to find out which area is the best one to open a Japanese restaurant.

Methodology

First, we will need to extract the data from the data sources:



List of postal codes of Canada: M

From Wikipedia, the free encyclopedia

This is a list of postal codes in Canada where the first letter is M. Postal codes beginning with M (except M0R and M7R) are located within the city of Toronto in the province of Ontario. Only the first three characters are listed, corresponding to the Forward Sortation Area. Canada Post provides a free postal code look-up tool on its website,[1] via its applications for such smartphones as the iPhone and BlackBerry,[2] and sells hard-copy directories and CD-ROMs. Many vendors also sell validation tools, which allow customers to properly match addresses and postal codes. Hard-copy directories can also be consulted in all post offices, and some libraries.

Toronto - 103 FSAs [edit]

Note: There are no rural FSAs in Toronto, hence no postal codes should start with M0. However, a handful of individual special-purpose codes in the M0R FSA are assigned to Gateway Commercial Returns, 4567 Dixie Rd, Mississauga as a merchandise returns label for freepost returns to high-volume vendors such as Amazon and the Shopping Channel.[3]

M1A Not assigned	M2A 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 Queen's Park (Ontario Provincial Government)	M8A Not assigned	M9A Etobicoke (Islington Avenue)
M1B Scarborough (Malvern / Rouge)	M2B Not assigned	M3B North York (Don Mills) North	M4B East York (Parkview Hill / Woodbine Gardens)	M5B Downtown Toronto (Garden District, Ryerson)	M6B North York (Glencairn)	M7B Not assigned	M8B Not assigned	M9B Etobicoke (West Deane Park / Princess Gardens / Martin Grove / Islington / Cloverdale)
M1C Scarborough (Rouge Hill / Port Union / Highland Creek)	M2C Not assigned	M3C North York (Don Mills) South (Flemingdon Park)	M4C East York (Woodbine Heights)	M5C Downtown Toronto (St. James Town)	M6C York (Humewood-Cedarvale)	M7C Not assigned	M8C Not assigned	M9C Etobicoke (Eringate / Bloordale Gardens / Old Burnhamthorpe / Markland Wood)
M1E Scarborough (Guildwood / Morningside / West Hill)	M2E Not assigned	M3E Not assigned	M4E East Toronto (The Beaches)	M5E Downtown Toronto (Berczy Park)	M6E York (Caledonia-Fairbanks)	M7E Not assigned	M8E Not assigned	M9E Not assigned
M1G Scarborough (Woburn)	M2G Not assigned	M3G Not assigned	M4G East York (Leaside)	M5G Downtown Toronto (Central Bay Street)	M6G Downtown Toronto (Christie)	M7G Not assigned	M8G Not assigned	M9G Not assigned
M1H	M2H	M3H	M4H	M5H	M6H	M7H	M8H	M9H

Fig.1: Toronto Neighborhoods via Wikipedia

The Wikipedia site (https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M) shown above, provided almost all the information about the neighborhoods. It included the postal code, borough and the name of the neighborhoods present in Toronto.

The data from the above site was scraped using BeautifulSoup and put into a dataframe as shown below:

```
[6]: df_Toronto.head(10)
```

[6]:

	PostalCode	Borough	Neighborhood
0	M3A	North York	Parkwoods
1	M4A	North York	Victoria Village
2	M5A	Downtown Toronto	Regent Park, Harbourfront
3	M6A	North York	Lawrence Manor, Lawrence Heights
4	M7A	Queen's Park	Ontario Provincial Government
5	M9A	Etobicoke	Islington Avenue
6	M1B	Scarborough	Malvern, Rouge
7	M3B	North York	Don Mills North
8	M4B	East York	Parkview Hill, Woodbine Gardens
9	M5B	Downtown Toronto	Garden District, Ryerson

Fig.2: Dataframe containing the information about neighborhoods

The second source of data (https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDeveloperSkillsNetwork-DS0701EN-SkillsNetwork/labs_v1/Geospatial_Coordinates.csv) provided us with the Geographical coordinates of the neighborhoods with the respective Postal Codes. The file was in CSV format, so attaching it to a Pandas data frame was simple (shown in figure 3).

```
df_geo_coor.head()
```

[7]:

	Postal Code	Latitude	Longitude
0	M1B	43.806686	-79.194353
1	M1C	43.784535	-79.160497
2	M1E	43.763573	-79.188711
3	M1G	43.770992	-79.216917
4	M1H	43.773136	-79.239476

```
[ ]:
```

Fig. 3: Geographical coordinates of each postal code

Both dataframes obtained were merged into one, so that each neighborhood has its geographical coordinates assigned. After that, the postal code column was dropped as we don’t need it when analyzing.

```
df_toronto.head()
```

[8]:

	PostalCode	Borough	Neighborhood	Latitude	Longitude
0	M3A	North York	Parkwoods	43.753259	-79.329656
1	M4A	North York	Victoria Village	43.725882	-79.315572
2	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636
3	M6A	North York	Lawrence Manor, Lawrence Heights	43.718518	-79.464763
4	M7A	Queen's Park	Ontario Provincial Government	43.662301	-79.389494

Fig. 4: Dataframes merged

```
df_toronto.head()
```

[9]:

	Borough	Neighborhood	Latitude	Longitude
0	North York	Parkwoods	43.753259	-79.329656
1	North York	Victoria Village	43.725882	-79.315572
2	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636
3	North York	Lawrence Manor, Lawrence Heights	43.718518	-79.464763
4	Queen's Park	Ontario Provincial Government	43.662301	-79.389494

Fig. 5: Postal Code column dropped.

I used Nominatim to get the latitude and longitude of Toronto. Then I use Folium to get the map of Toronto and its neighborhoods.

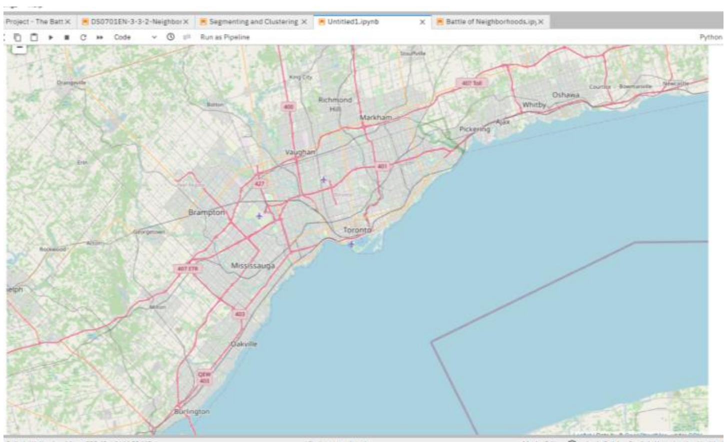


Fig. 6: Map of Toronto

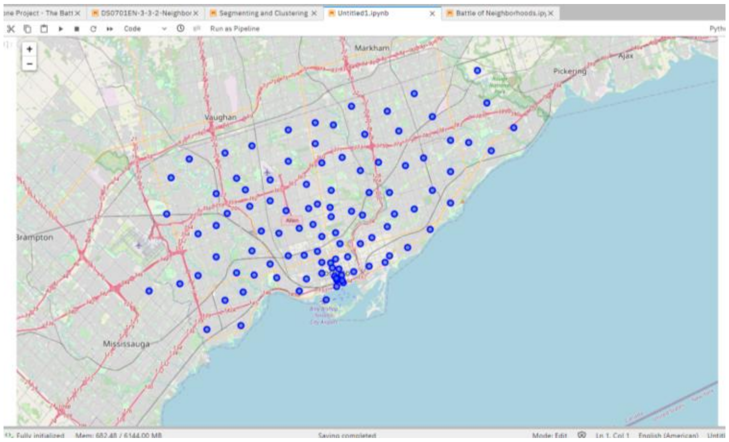


Fig. 6: Toronto and its neighborhoods

The retrieval of the location, name, category and food type about the restaurants in Toronto was collected through the Foursquare explore API. To obtain the data, it was required to make an account where it would provide a ‘Secret Key’ as well as a ‘Client ID’ which would allow me to pull any data.

```
restaurants_toronto
```

[18]:

	Neighborhood	ID	Name	Latitude	Longitude	Category	Food
0	Parkwoods	4b8991cbf964a520814232e3	Allwyn's Bakery	43.759840	-79.324719	Caribbean Restaurant	Caribbean
1	Parkwoods	4b0aed06f964a520202a23e3	Island Foods	43.745866	-79.346035	Caribbean Restaurant	Caribbean
2	Parkwoods	4b149ea4f964a52029a523e3	Darband Restaurant	43.755194	-79.348498	Middle Eastern Restaurant	Middle Eastern
3	Parkwoods	519e5c39498eb25c945e98c8	VIA CIBO | italian streetfood	43.754067	-79.357951	Italian Restaurant	Italian
4	Parkwoods	54b55e81498e6b087da5f439	Me Va Me Kitchen Express	43.754957	-79.351894	Mediterranean Restaurant	Mediterranean
...	...	...	...	...	...	...	...
878	Enclave of M4L	4bc37b7c920eb713b4851d2c	Ruby WatchCo.	43.659149	-79.349170	Restaurant	Restaurant
879	Old Mill South, King's Mill Park, Sunnylea, Hu...	543b1213498e97d631e83226	U-Know Sushi	43.630326	-79.485386	Sushi Restaurant	Sushi
880	Old Mill South, King's Mill Park, Sunnylea, Hu...	4b527767f964a520a37e27e3	Eden Trattoria	43.627035	-79.476776	Italian Restaurant	Italian
881	Old Mill South, King's Mill Park, Sunnylea, Hu...	4b15aef7f964a52074b223e3	Asa Sushi	43.649902	-79.484611	Sushi Restaurant	Sushi
882	Old Mill South, King's Mill Park, Sunnylea, Hu...	4bd2e88077b29c74e9838f82	Bloom Restaurant	43.650307	-79.479836	Latin American Restaurant	Latin American

883 rows × 7 columns

Fig. 7: Restaurants in Toronto

We then created a map using folium and color coded the restaurants: the Japanese ones are marked with red. Also, we can see that there are 883 restaurants in Toronto.

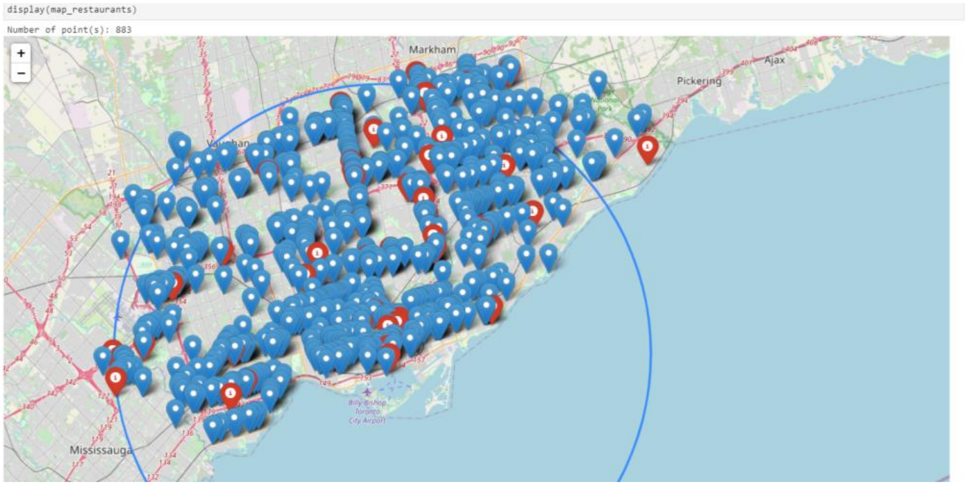


Fig. 8: Map of restaurants in Toronto

We filtered the Japanese restaurants and then created a map with all of them in the entire Toronto. We can see that there are 52 Japanese Restaurants in Toronto.

Leaflet | Data by © OpenStreetMap, under ODbL.

```
[20]: jap_restaurants_toronto = restaurants_toronto[restaurants_toronto['Category'].str.contains('Japanese')]
      jap_restaurants_toronto = jap_restaurants_toronto.drop_duplicates(subset=['ID']).reset_index(drop = True) ## Remove items ID duplicates and reset index
      jap_restaurants_toronto
```

[20]:

	Neighborhood	ID	Name	Latitude	Longitude	Category	Food
0	Parkwoods	53bafb4f498eb927faa3cd9e	Matsuda Japanese Cuisine & Teppanyaki	43.745494	-79.345821	Japanese Restaurant	Japanese
1	Parkwoods	4b78331ef964a520b9bb2ee3	Samjung Japanese Restaurant	43.755251	-79.360323	Japanese Restaurant	Japanese
2	Parkwoods	4b639286f964a5203c842ae3	Gonoe Sushi	43.745737	-79.345991	Japanese Restaurant	Japanese
3	Victoria Village	4b36ac26f964a520623a25e3	Kaiseki Yu-Zen Hashimoto	43.720931	-79.337185	Japanese Restaurant	Japanese
4	Victoria Village	4b7441c4f964a5205cd02de3	Teriyaki Experience	43.727006	-79.292946	Japanese Restaurant	Japanese
5	Regent Park, Harbourfront	574ad72238fa943556d93b8e	Gyu-Kaku Japanese BBQ	43.651422	-79.375047	Japanese Restaurant	Japanese
6	Regent Park, Harbourfront	50e8b903e4b04fa56858ecfe	Zakkushi ざっ串	43.663618	-79.370670	Japanese Restaurant	Japanese
7	Regent Park, Harbourfront	56201ed4498e7f700c462170	Miku	43.641374	-79.377531	Japanese Restaurant	Japanese
8	Regent Park, Harbourfront	5a4fdf56772fbc5e9fa73c7f	Chotto Matte	43.646473	-79.378782	Japanese Restaurant	Japanese
9	Regent Park, Harbourfront	4ae73054f964a5203ca921e3	Ki Modern Japanese + Bar	43.647223	-79.379374	Japanese Restaurant	Japanese
10	Lawrence Manor, Lawrence Heights	5547aad8498e00f3f580ea93	Sushi Zuki	43.694494	-79.456141	Japanese Restaurant	Japanese
11	Lawrence Manor, Lawrence Heights	4fca732be4b07c8e3e198c27	Miyako Sushi Restaurant	43.709111	-79.443930	Japanese Restaurant	Japanese
12	Ontario Provincial Government	4b2bd898f964a52042bc24e3	Kinka Izakaya Original	43.660596	-79.378891	Japanese Restaurant	Japanese
13	Ontario Provincial Government	5362c366498e602fbe1db395	Yasu	43.662837	-79.403217	Japanese Restaurant	Japanese
14	Islington Avenue	4b057d9af964a520c25822e3	Gojima	43.645094	-79.521672	Japanese Restaurant	Japanese
15	Parkview Hill, Woodbine Gardens	51df3790498e737986e114bb	Itamae AYCE Sushi	43.713998	-79.335484	Japanese Restaurant	Japanese
16	Rouge Hill, Port Union, Highland Creek	513fd3dae4b0e4688c9cf6a8	Azumi Teppanyaki and Sushi	43.779910	-79.138464	Japanese Restaurant	Japanese
17	Eringate, Bloordale Gardens, Old Burnhamthorpe...	4b12efd8f964a5202e9123e3	Prince Japanese Steakhouse	43.648952	-79.605052	Japanese Restaurant	Japanese
18	The Beaches	587bae87286804584acd2a7a	Isabella's Boutique Restaurant	43.673767	-79.282703	Japanese Restaurant	Japanese
19	Leaside	589f58fbd0bb3e25a8e5a88a	Rolltation	43.707517	-79.398456	Japanese Restaurant	Japanese
20	Christie	56418c69498ec00425302f6a	Imanishi Japanese Kitchen	43.649546	-79.426551	Japanese Restaurant	Japanese
21	Cedarbrae	4aeb1ce4f964a52086be21e3	Ni-Ji	43.767625	-79.271399	Japanese Restaurant	Japanese

Idle Fully initialized Mem: 751.57 / 6144.00 MB Saving completed Mode: Edit Ln 1, Col 1 English (American) Uni

Fig. 9: Japanese Restaurants in Toronto

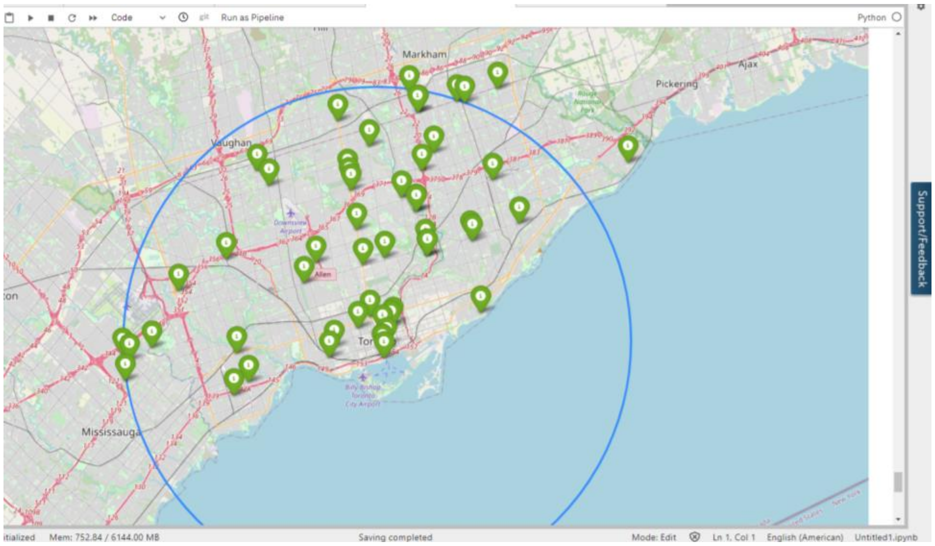


Fig. 10: Map of Japanese Restaurants in Toronto

To analyze the data we performed a technique in which Categorical Data is transformed into Numerical Data for Machine Learning algorithms. This technique is called One hot encoding. For each of the neighborhoods, individual restaurants were turned into the frequency at how many of those restaurants were located in each neighborhood.

```
neighborhood_grouped = neighborhood_onehot.groupby('Neighborhood').mean().reset_index()
neighborhood_grouped
```

[27]:

	Neighborhood	Afghan	African	American	Asian	Brazilian	Cajun / Creole	Cantonese	Caribbean	Caucasian	...	Szechuan	Taiwanese	Tapas	Thai	Theme Restaurant	Tibetan	Turkish	Vegetarian / Vegan	Vietnamese
0	Agincourt	0.0	0.0	0.000000	0.071429	0.0	0.0	0.071429	0.071429	0.0	...	0.0	0.0	0.0	0.000000	0.0	0.0	0.000000	0.0	0.071429
1	Alderwood, Long Branch	0.0	0.0	0.066667	0.000000	0.0	0.0	0.000000	0.000000	0.0	...	0.0	0.0	0.0	0.000000	0.0	0.0	0.000000	0.0	0.000000
2	Bathurst Manor, Wilson Heights, Downsview North	0.0	0.0	0.000000	0.038462	0.0	0.0	0.000000	0.000000	0.0	...	0.0	0.0	0.0	0.076923	0.0	0.0	0.000000	0.0	0.000000
3	Bayview Village	0.0	0.0	0.000000	0.000000	0.0	0.0	0.000000	0.000000	0.0	...	0.0	0.0	0.0	0.058824	0.0	0.0	0.000000	0.0	0.000000
4	Bedford Park, Lawrence Manor East	0.0	0.0	0.000000	0.111111	0.0	0.0	0.000000	0.000000	0.0	...	0.0	0.0	0.0	0.000000	0.0	0.0	0.111111	0.0	0.000000
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
78	Willowdale West	0.0	0.0	0.000000	0.000000	0.0	0.0	0.000000	0.000000	0.0	...	0.0	0.0	0.0	0.000000	0.0	0.0	0.000000	0.0	0.100000
79	Willowdale, Newtonbrook	0.0	0.0	0.000000	0.000000	0.0	0.0	0.040000	0.000000	0.0	...	0.0	0.0	0.0	0.000000	0.0	0.0	0.000000	0.0	0.000000
80	Woburn	0.0	0.0	0.071429	0.000000	0.0	0.0	0.000000	0.071429	0.0	...	0.0	0.0	0.0	0.071429	0.0	0.0	0.000000	0.0	0.000000
81	Woodbine Heights	0.0	0.0	0.117647	0.117647	0.0	0.0	0.000000	0.058824	0.0	...	0.0	0.0	0.0	0.058824	0.0	0.0	0.058824	0.0	0.000000
82	York Mills, Silver Hills	0.0	0.0	0.000000	0.000000	0.0	0.0	0.000000	0.000000	0.0	...	0.0	0.0	0.0	0.000000	0.0	0.0	0.000000	0.0	0.000000

83 rows × 62 columns

Fig. 11: One hot encoding

We then searched for the most commons restaurants for each neighborhood and overall. We can see that overall the Japanese Restaurants are on the 5^{th} position.

```
neighborhoods_venues_sorted
```

	Neighborhood	1st Most Common Restaurant	2nd Most Common Restaurant	3rd Most Common Restaurant	4th Most Common Restaurant	5th Most Common Restaurant
0	Agincourt	Chinese	Indian	Fast Food	Cantonese	Restaurant
1	Alderwood, Long Branch	Fast Food	Seafood	Restaurant	Chinese	Middle Eastern
2	Bathurst Manor, Wilson Heights, Downsview North	Restaurant	Sushi	Middle Eastern	Japanese	Ramen
3	Bayview Village	Korean	Sushi	Ramen	Middle Eastern	Seafood
4	Bedford Park, Lawrence Manor East	Sushi	Fast Food	Turkish	Asian	Japanese
...	...	...	...	...	...	...
78	Willowdale West	Korean	Middle Eastern	Sushi	Vietnamese	Restaurant
79	Willowdale, Newtonbrook	Korean	Middle Eastern	Sushi	Fast Food	Indian
80	Woburn	Indian	Fast Food	Restaurant	Caribbean	Hakka
81	Woodbine Heights	Indian	American	Asian	Caribbean	Fast Food
82	York Mills, Silver Hills	Restaurant	Italian	Chinese	Indian	Middle Eastern

83 rows × 6 columns

Fig. 12: 5 most common restaurants for each neighborhood

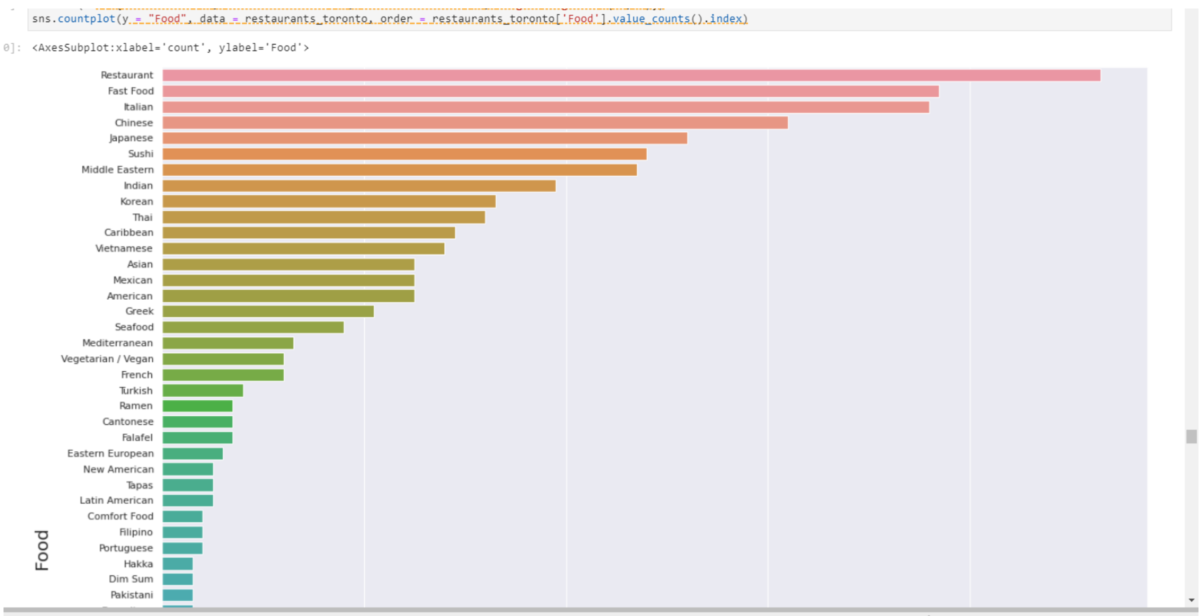


Fig. 13: Most common restaurants in Toronto

We also wanted to cluster the neighborhoods to see which are has the most Japanese Restaurants. To do this we used K-Means clustering. To get our optimum K value that was neither overfitting or underfitting the model, we used the Elbow Point Technique. In this technique we ran a test with different number of K values and measured the accuracy and then chose the best K value. The best K value is chosen at the point in which the line has a sharpest turn. In our case we had the Elbow Point at K = 4. That means we will have a total of 4 clusters.

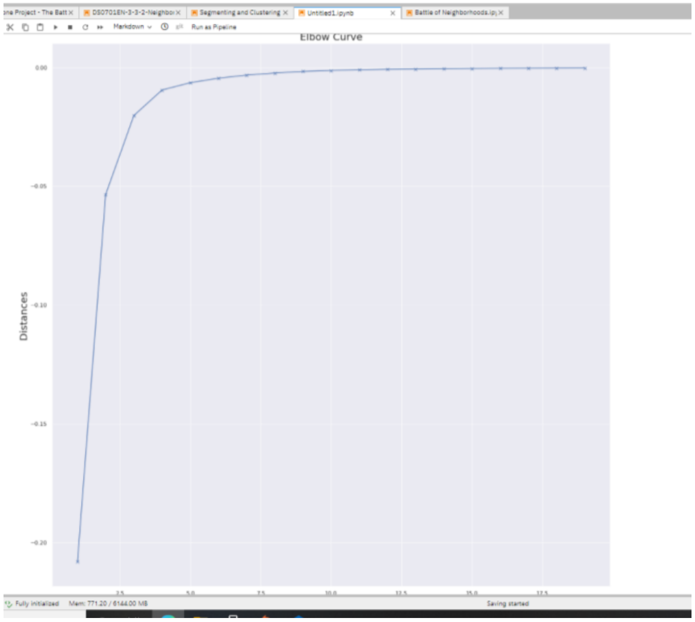
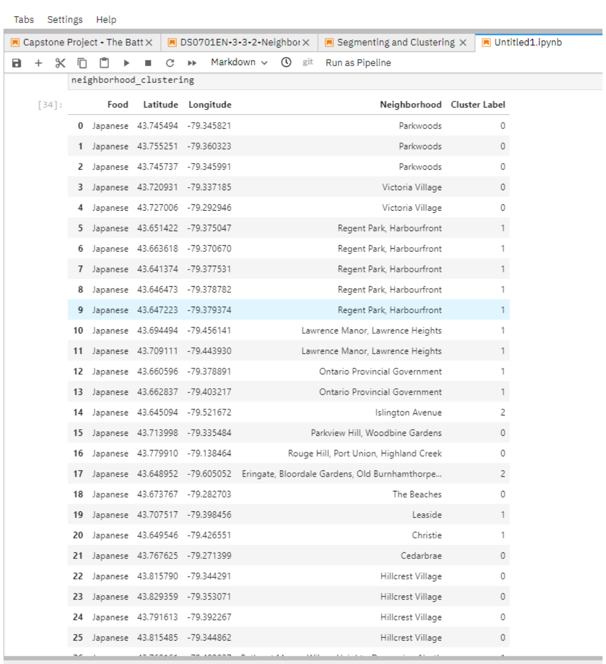


Fig. 14: Finding the K vs Error Values

We see that the Elbow is at K=3. Moreover, in K-Means clustering, objects that are similar based on a certain variable are put into the same cluster. Neighborhoods that had similar mean frequency of Japanese Restaurants were divided into 3 clusters. Each of these clusters were labelled from 0 to 3 as the indexing of labels begin with 0 instead of 1.



```
neighborhood_clustering
```

	Food	Latitude	Longitude	Neighborhood	Cluster Label
0	Japanese	43.745494	-79.345821	Parkwoods	0
1	Japanese	43.755251	-79.360323	Parkwoods	0
2	Japanese	43.745737	-79.345991	Parkwoods	0
3	Japanese	43.720931	-79.337185	Victoria Village	0
4	Japanese	43.727006	-79.292946	Victoria Village	0
5	Japanese	43.651422	-79.375047	Regent Park, Harbourfront	1
6	Japanese	43.663618	-79.370670	Regent Park, Harbourfront	1
7	Japanese	43.641374	-79.377531	Regent Park, Harbourfront	1
8	Japanese	43.646473	-79.378782	Regent Park, Harbourfront	1
9	Japanese	43.647223	-79.379374	Regent Park, Harbourfront	1
10	Japanese	43.694494	-79.456141	Lawrence Manor, Lawrence Heights	1
11	Japanese	43.709111	-79.443930	Lawrence Manor, Lawrence Heights	1
12	Japanese	43.660596	-79.378891	Ontario Provincial Government	1
13	Japanese	43.662837	-79.403217	Ontario Provincial Government	1
14	Japanese	43.645094	-79.521672	Islington Avenue	2
15	Japanese	43.713998	-79.335484	Parkview Hill, Woodbine Gardens	0
16	Japanese	43.779910	-79.138464	Rouge Hill, Port Union, Highland Creek	0
17	Japanese	43.648952	-79.605052	Eringate, Bloordale Gardens, Old Burnhamthorpe...	2
18	Japanese	43.673767	-79.282703	The Beaches	0
19	Japanese	43.707517	-79.398456	Leaside	1
20	Japanese	43.649546	-79.426551	Christie	1
21	Japanese	43.767625	-79.271399	Cedarbrae	0
22	Japanese	43.815790	-79.344291	Hillcrest Village	0
23	Japanese	43.829359	-79.353071	Hillcrest Village	0
24	Japanese	43.791613	-79.392267	Hillcrest Village	0
25	Japanese	43.815485	-79.344862	Hillcrest Village	0

Fig. 15: Appropriate Cluster Labels were added

After, we merged the venue data with the table above creating a new table which would be the basis for analyzing new opportunities for opening a new Japanese Restaurant in Toronto. Then we created a map using the Folium package in Python and each neighborhood was colored based on the cluster label. For example, cluster 1 was purple and cluster 2 was green.

```
map_clusters
```

[39]:

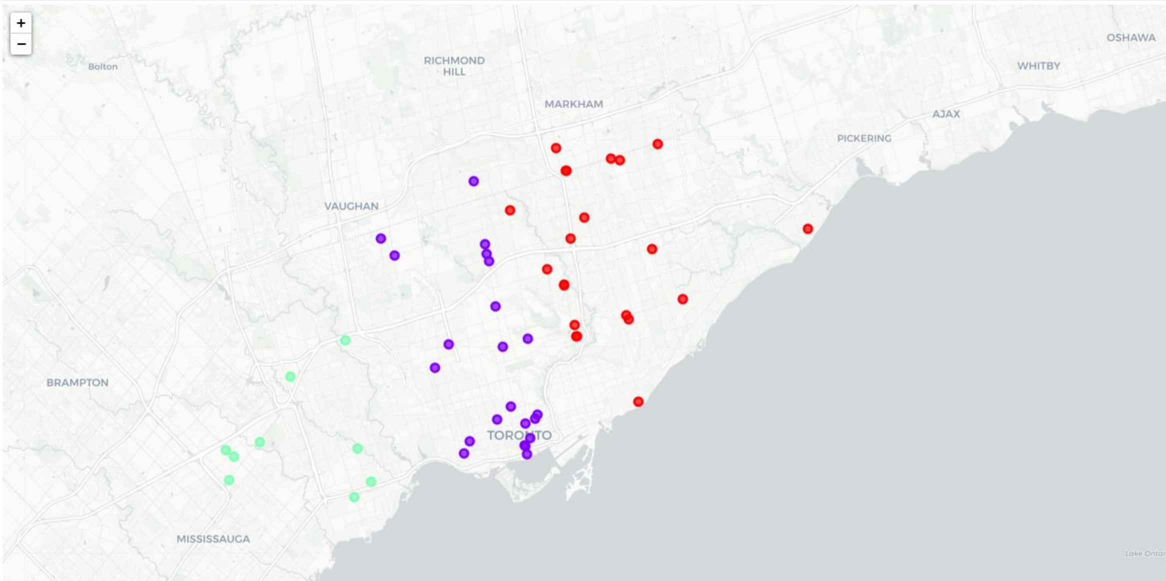


Fig. 16: Map with different Clusters

Cluster 1:

[40]:
```
## Cluster 1 in neighborhood_merged
neighborhood_cluster1 = neighborhood_clustering[neighborhood_clustering['Cluster Label'] == 0]
neighborhood_cluster1
```

[40]:

	Food	Latitude	Longitude	Neighborhood	Cluster Label
0	Japanese	43.745494	-79.345821	Parkwoods	0
1	Japanese	43.823043	-79.306446	Milliken, Agincourt North, Steeles East, L'Amo...	0
2	Japanese	43.831832	-79.266304	Milliken, Agincourt North, Steeles East, L'Amo...	0
3	Japanese	43.822372	-79.298905	Milliken, Agincourt North, Steeles East, L'Amo...	0
4	Japanese	43.724299	-79.290987	Kennedy Park, Ionview, East Birchmount Park	0
5	Japanese	43.774171	-79.340401	Fairview, Henry Farm, Oriole	0
6	Japanese	43.787145	-79.328940	Fairview, Henry Farm, Oriole	0
7	Japanese	43.737003	-79.244750	Scarborough Village	0
8	Japanese	43.714044	-79.335000	Thorncliffe Park	0
9	Japanese	43.791613	-79.392267	Hillcrest Village	0
10	Japanese	43.829359	-79.353071	Hillcrest Village	0
11	Japanese	43.815790	-79.344291	Hillcrest Village	0
12	Japanese	43.767625	-79.271399	Cedarbrae	0
13	Japanese	43.673767	-79.282703	The Beaches	0
14	Japanese	43.779910	-79.138464	Rouge Hill, Port Union, Highland Creek	0
15	Japanese	43.815485	-79.344862	Hillcrest Village	0
16	Japanese	43.755251	-79.360323	Parkwoods	0
17	Japanese	43.745737	-79.345991	Parkwoods	0
18	Japanese	43.720931	-79.337185	Victoria Village	0
19	Japanese	43.727006	-79.292946	Victoria Village	0
20	Japanese	43.713998	-79.335484	Parkview Hill, Woodbine Gardens	0

Cluster 2:

[41]:
```
## Cluster 2 in neighborhood_merged
neighborhood_cluster2 = neighborhood_clustering[neighborhood_clustering['Cluster Label'] == 1]
neighborhood_cluster2
```

[41]:

	Food	Latitude	Longitude	Neighborhood	Cluster Label
21	Japanese	43.641846	-79.431086	Little Portugal, Trinity	1
22	Japanese	43.670683	-79.391056	Moore Park, Summerhill East	1
23	Japanese	43.712637	-79.376914	Lawrence Park	1
24	Japanese	43.774224	-79.501720	Downsview Northwest	1
25	Japanese	43.763523	-79.490136	Downsview Northwest	1
26	Japanese	43.732562	-79.404147	Bedford Park, Lawrence Manor East	1
27	Japanese	43.809168	-79.423015	Willowdale, Newtonbrook	1
28	Japanese	43.651422	-79.375047	Regent Park, Harbourfront	1
29	Japanese	43.663618	-79.370670	Regent Park, Harbourfront	1
30	Japanese	43.665895	-79.368415	Rosedale	1
31	Japanese	43.646473	-79.378782	Regent Park, Harbourfront	1
32	Japanese	43.764412	-79.411880	Bathurst Manor, Wilson Heights, Downsview North	1
33	Japanese	43.770806	-79.413138	Bathurst Manor, Wilson Heights, Downsview North	1
34	Japanese	43.760161	-79.409827	Bathurst Manor, Wilson Heights, Downsview North	1
35	Japanese	43.647223	-79.379374	Regent Park, Harbourfront	1
36	Japanese	43.694494	-79.456141	Lawrence Manor, Lawrence Heights	1
37	Japanese	43.709111	-79.443930	Lawrence Manor, Lawrence Heights	1
38	Japanese	43.660596	-79.378891	Ontario Provincial Government	1
39	Japanese	43.641374	-79.377531	Regent Park, Harbourfront	1
40	Japanese	43.662837	-79.403217	Ontario Provincial Government	1
41	Japanese	43.707517	-79.398456	Leaside	1
42	Japanese	43.649546	-79.426551	Christie	1

[42]:
```
## Cluster 3 in neighborhood_merged
```

Cluster 3:

42	Japanese	43.649546	-79.426551	Christie	1

[42]:
```
## Cluster 3 in neighborhood_merged
neighborhood_cluster3 = neighborhood_clustering[neighborhood_clustering['Cluster Label'] =
neighborhood_cluster3
```

[42]:

	Food	Latitude	Longitude	Neighborhood	Cluster Label
43	Japanese	43.615389	-79.524415	New Toronto, Mimico South, Humber Bay Shores	2
44	Japanese	43.648952	-79.605052	Eringate, Bloordale Gardens, Old Burnhamthorpe...	2
45	Japanese	43.640005	-79.626960	Enclave of L4W	2
46	Japanese	43.625661	-79.630920	Enclave of L4W	2
47	Japanese	43.644181	-79.633899	Enclave of L4W	2
48	Japanese	43.711507	-79.531797	Humberlea, Emery	2
49	Japanese	43.624728	-79.509904	New Toronto, Mimico South, Humber Bay Shores	2
50	Japanese	43.689193	-79.578441	Kingsview Village, St. Phillips, Martin Grove ...	2
51	Japanese	43.645094	-79.521672	Islington Avenue	2

Considering the above, the ordering of the average Japanese Restaurant in each cluster goes as follows:

1. Cluster 1 (red) = Cluster 2 (purple): with 21 Japanese restaurants those 2 areas are on the same position
2. Cluster 3 (green): with only 8 Japanese restaurants

Discussion

Most of the Japanese Restaurants are in cluster 1 and cluster 2 represented by the red and purple clusters. The neighborhoods in these areas that have the highest average of Japanese Restaurants are Hillcrest Village and Regent Park. We can see that in the cluster 3 there are significantly fewer Japanese restaurants. However, in this area the most restaurants are in Enclave of L4W. Looking at the nearby venues, the optimum place to put a new Japanese restaurant is in the cluster 3 area as there are many Neighborhoods in the area but little to no Japanese restaurants. Some of the drawbacks of this analysis are – the clustering is completely based on data obtained from Foursquare API. Also, the analysis does not take into consideration of the Japanese population across neighborhoods as this can play a huge factor while choosing which place to open a new Japanese restaurant. This concludes the optimal findings for this project and recommends the entrepreneur to open an authentic Japanese restaurant in these locations with little to no competition.

Conclusion

In conclusion, to end off this project, we had an opportunity on a business problem, and it was tackled in way that it was similar to how a genuine data scientist would do. We utilized numerous Python libraries to fetch the information , to control the content and to break down and visualize those datasets. We have utilized Foursquare API to investigate the settings in neighborhoods of Toronto, get great measure of data from Wikipedia which we scraped with the Beautifulsoup Web scraping Library. We also visualized utilizing different plots present in seaborn and matplotlib libraries. Similarly, we applied AI strategy to anticipate the error given the information and utilized Folium to picture it on a map.

Places that have room for improvement or certain drawbacks gives us that this project can be additionally improved with the assistance of more information and distinctive Machine Learning strategies. Additionally, we can utilize this venture to investigate any situation, for example, opening an alternate cuisine or opening of a Movie Theater and so forth. Ideally, this task acts as an initial direction to tackle more complex real-life problems using data-science.