

Mexican Restaurant in Toronto

Downtown Area

The Battle of Neighbourhood Capstone Project

A full report consisting of all of the following components

~Introduction where you discuss the business problem and who would be interested in this project.

~Data where you describe the data that will be used to solve the problem and the source of the data.

~Methodology section which represents the main component of the report where you discuss and describe any exploratory data analysis that you did, any inferential statistical testing that you performed, if any, and what machine learnings were used and why.

~Results section where you discuss the results.

~Discussion section where you discuss any observations you noted and any recommendations you can make based on the results.

~Conclusion section where you conclude the report.

Part I

Description of the problem

Toronto's competitive restaurant scene is an impressive one thanks to the various chefs and restaurateurs who constantly bring people new places and cuisines to try. But, where is the best place to open up a Mexican eatery in the Greater Toronto Area?

Over the next 50 years, the city is projected to grow at a thrilling rate. In 50 years, the GTA's population will double. The city's diverse population, currently at 2.9 million, will increase to almost 5 million. Nearly 10 per cent of new Torontonians will cram themselves into the core, a sliver that accounts for only three per cent of the city's land mass.

As a result - all things being equal and acknowledging the fact that opening a restaurant business requires deep pockets and lots of hard work, it would seem to be a good choice to open up a Mexican eatery in the core of the city. The question to be answered then is, which neighborhoods in Downtown Toronto would represent the best option.

Discussion of the background

My client, a wealthy and successful restaurateur from Mexico is eager to expand business operations into Toronto. They want to create an authentic Mexican eatery that will serve and offer the full richness of Mexican culture and cuisine to the people of Downtown Toronto.

Since Downtown Toronto is very competitive, my client needs insight from data in order to decide in which neighborhood to establish this authentic Mexican eatery.

Part II

Description of the data

This project will utilize publicly available data from Wikipedia and Foursquare.

Specifically, all Toronto neighborhood details along with their postal codes are available here: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

The focus of this project will be the Downtown Toronto neighborhoods that will be extracted and analyzed accordingly.

The Foursquare API will be utilized to obtain the geographical location data for Downtown Toronto, and data will be used to explore the restaurant venues in the neighbourhoods. The restaurants will provide the categories needed for the analysis and these will be used to determine the viability of the selected locations for the restaurant.

How data will be used to solve the problem

The data from Wikipedia and Foursquare will be explored and analyzed by considering the restaurant venues in Downtown Toronto. The restaurants from the core of the city will be reviewed in terms of the types or categories of restaurants within a specific radius.

The data will be utilized to come up with a frequency analysis for a Mexican eatery in Downtown Toronto, and to come up with the best choices of neighborhoods for my client.

1.Importing all the libraries

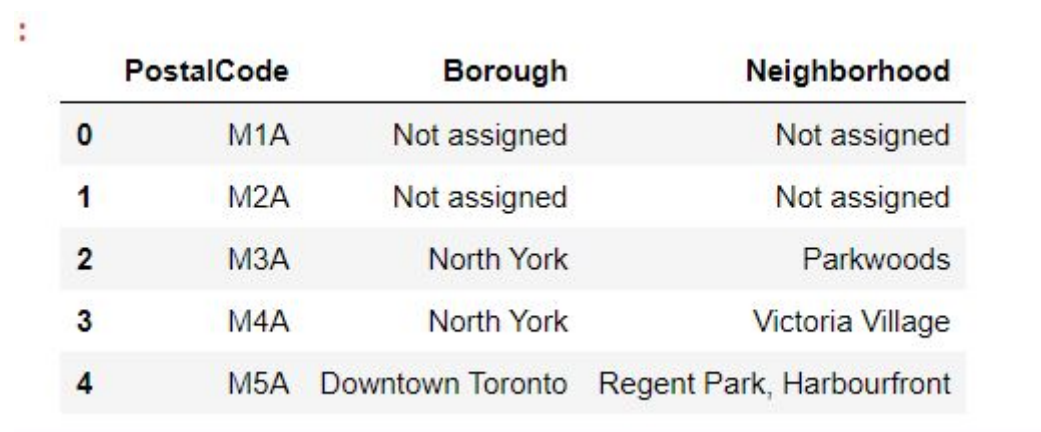
I downloaded numpy for library to handle data in a vectorized manner, pandas for data analysis, Json library to handle JSON files, Matplotlib for plotting

2. Extracting dataset from the specified url

All the Toronto neighborhood details along with their postal codes are available here: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

3. Data Wrangling & Cleaning

I First discarded the reference/citation links, cleaned the header row of the sort icons, defined the dataframe columns as postal Code, Borough 'm' Neighbourhood, Instantiated & populated the dataframe & examined the resulting dataframe.



	PostalCode	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	Regent Park, Harbourfront

Then I Processed the cells that have an assigned borough.I had to Ignore cells with a borough that is not assigned. The neighborhood had to be same as the borough if a cell has a borough but with a 'Not' assigned neighborhood.

More than one neighborhood can exist in one postal code area so i Combined the rows into one row with the neighborhoods separated with a comma.

	PostalCode	Borough	Neighborhood
0	M1B	Scarborough	Malvern, Rouge
1	M1C	Scarborough	Rouge Hill, Port Union, Highland Creek
2	M1E	Scarborough	Guildwood, Morningside, West Hill
3	M1G	Scarborough	Woburn
4	M1H	Scarborough	Cedarbrae

Finally I Made a copy of the dataframe.

```

The number of rows in dataframe: 103
: (103, 3)

```

4. Importing Geospacial data for Latitude and Longitude usage in the data

https://cocl.us/Geospacial_data

	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

5. Merging it to the original dataset

	PostalCode	Borough	Neighborhood	Latitude	Longitude
0	M1B	Scarborough	Malvern, Rouge	43.806686	-79.194353
1	M1C	Scarborough	Rouge Hill, Port Union, Highland Creek	43.784535	-79.160497
2	M1E	Scarborough	Guildwood, Morningside, West Hill	43.763573	-79.188711
3	M1G	Scarborough	Woburn	43.770992	-79.216917
4	M1H	Scarborough	Cedarbrae	43.773136	-79.239476

I Merged the original df_copy with geodata & Checked how many boroughs and neighborhoods are there.

For map i Imported Folium visualization library Segmented and Clustered it by Downtown Toronto.

	PostalCode	Borough	Neighborhood	Latitude	Longitude
0	M4W	Downtown Toronto	Rosedale	43.679563	-79.377529
1	M4X	Downtown Toronto	St. James Town, Cabbagetown	43.667967	-79.367675
2	M4Y	Downtown Toronto	Church and Wellesley	43.665860	-79.383160
3	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636
4	M5B	Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937

6. Preparing for Map view

Converted an address into latitude and longitude values by using geopy library for the latitude and longitude values of Toronto.

Then created a map of Downtown Toronto using Latitude and Longitude values and added markers to the map.



7. Foursquare API usage

Next, I started utilizing the Foursquare API to explore the neighborhoods and segmented them. Used the Foursquare ID, Foursquare Secret and the Foursquare API version.

I got the neighborhood name and the neighborhoods' latitude and longitude values.

8. Top 100 venues within 500 meters

Got the top 100 venues that are in Rosedale within a radius of 500 meters from this link below

https://api.foursquare.com/v2/venues/explore?client_id=RDZDO5MSITL4N20HVKR2WPZ1RHFP3JVEI1OZHSLTRJC1MYMX&client_secret=WB3GMUZTGTJCTXN0CSCJGG54F0UBXRE1IWC5WMDJ1KSCAM4X&v=20180605&ll=43.6056466,-79.50132070000001&radius=500&limit=100

Sent the GET request and examined the results.

All the information were in the items key. Borrowed the `get_category_type` function from the Foursquare lab.

Cleaned the json and structured it into a pandas dataframe, filtered the columns filtered the category for each row cleaned columns.



	name	categories	lat	lng
0	LCBO	Liquor Store	43.602281	-79.499302
1	New Toronto Fish & Chips	Restaurant	43.601849	-79.503281
2	Domino's Pizza	Pizza Place	43.601583	-79.500905
3	Delicia Bakery & Pastry	Bakery	43.601403	-79.503012
4	Lucky Dice Restaurant	Café	43.601392	-79.503056

Used the function from the lab to repeat the same process to all the neighborhoods in Downtown Toronto. Created the API request URL & made the GET request.

Returned only relevant information for each nearby venue. Ran the above function on each neighborhood and created a new dataframe called `tor_venues`.


```

Rosedale
St. James Town, Cabbagetown
Church and Wellesley
Regent Park, Harbourfront
Garden District, Ryerson
St. James Town
Berczy Park
Central Bay Street
Richmond, Adelaide, King
Harbourfront East, Union Station, Toronto Islands
Toronto Dominion Centre, Design Exchange
Commerce Court, Victoria Hotel
University of Toronto, Harbord
Kensington Market, Chinatown, Grange Park
CN Tower, King and Spadina, Railway Lands, Harbourfront West, Bathurst Quay, So
uth Niagara, Island airport
Stn A PO Boxes
First Canadian Place, Underground city
Christie
Queen's Park, Ontario Provincial Government

```

Finally Checked the size of the resulting dataframe & Found out how many unique categories can be curated from all the returned venues.

Also analyzed each neighborhood with one hot encoding and added neighborhood column back to dataframe.

```
tor_venues.tail()
```

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
1220	Queen's Park, Ontario Provincial Government	43.662301	-79.389494	SUDS	43.659880	-79.394712	Bar
1221	Queen's Park, Ontario Provincial Government	43.662301	-79.389494	Tim Hortons	43.658175	-79.390681	Coffee Shop
1222	Queen's Park, Ontario Provincial Government	43.662301	-79.389494	Convocation Hall	43.660828	-79.395245	College Auditorium
1223	Queen's Park, Ontario Provincial Government	43.662301	-79.389494	Tim Hortons	43.658906	-79.388696	Coffee Shop
1224	Queen's Park, Ontario Provincial Government	43.662301	-79.389494	Hart House Music Room	43.663758	-79.395027	Music Venue

The neighborhood column was moved to the first column, Examined the new dataframe size confirmed the new size, inspected temp, put that into a pandas dataframe & Used the function to sort the venues in descending order.

9. Venues to Decide and cluster through K Means for Restaurant

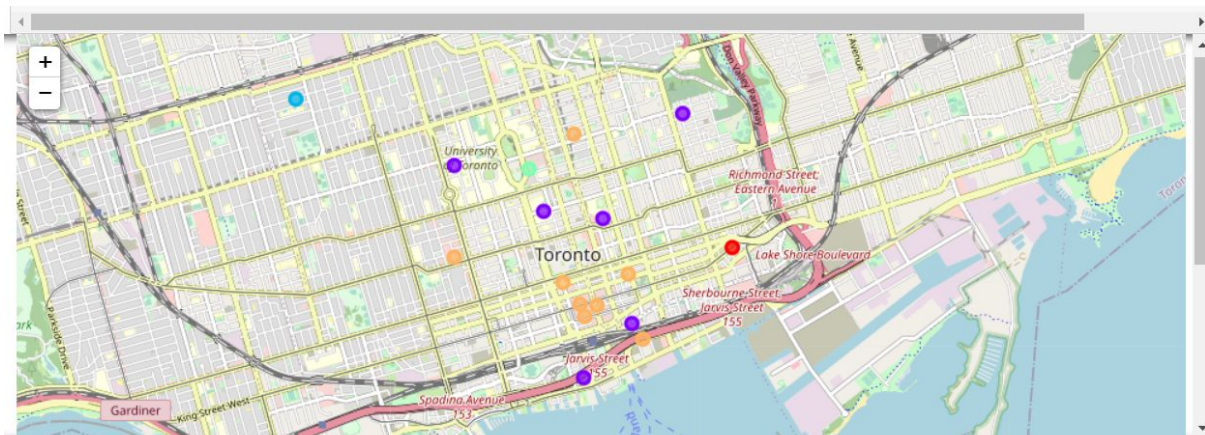
Created the new dataframe and displayed the top 15 venues for each neighborhood. Created columns according to number of top venues

Created a new dataframe, Ran k-means to cluster the neighborhood into 5 clusters and set number of clusters. Ran k-means clustering, checked cluster labels generated for each row in the dataframe, added clustering labels, merged to add latitude/longitude for each neighborhood & Ignored/dropped NaNs.

	PostalCode	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	
0	M4W	Downtown Toronto	Rosedale	43.679563	-79.377529	NaN	NaN	NaN	
1	M4X	Downtown Toronto	St. James Town, Cabbagetown	43.667967	-79.367675	1.0	Restaurant	Italian Restaurant	R
2	M4Y	Downtown Toronto	Church and Wellesley	43.665860	-79.383160	4.0	Sushi Restaurant	Japanese Restaurant	R
3	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636	0.0	Restaurant	French Restaurant	Vik R
4	M5B	Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937	1.0	Italian Restaurant	Japanese Restaurant	R

```
Number of rows after dropping NaNs: 17
Number of NaNs: PostalCode          0
Borough                             0
Neighborhood                         0
Latitude                            0
Longitude                           0
Cluster Labels                       0
1st Most Common Venue               0
2nd Most Common Venue               0
3rd Most Common Venue               0
4th Most Common Venue               0
5th Most Common Venue               0
6th Most Common Venue               0
7th Most Common Venue               0
8th Most Common Venue               0
9th Most Common Venue               0
10th Most Common Venue              0
11th Most Common Venue              0
12th Most Common Venue              0
13th Most Common Venue              0
14th Most Common Venue              0
15th Most Common Venue              0
dtype: int64
```

Visualized the Clusters, created the map by setting color scheme for the clusters. Finally, added markers to the map.



10. Best areas to open a Mexican Cuisines Restuarant in Down town Area

CLUSTER 1

```
# Cluster 1
tor_merged.loc[tor_merged['Cluster Labels'] == 0, tor_merged.columns[[2] + list(range(6, tor_merged.shape[1]))]]
```

	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	11th Most Common Venue	12th Most Common Venue	13th Most Common Venue
2	Regent Park, Harbourfront	Restaurant	French Restaurant	Vietnamese Restaurant	Doner Restaurant	German Restaurant	Filipino Restaurant	Fast Food Restaurant	Falafel Restaurant	Ethiopian Restaurant	Eastern European Restaurant	Dumpling Restaurant	Dim Sum Restaurant	Re

CLUSTER 2

```
# Cluster 2
tor_merged.loc[tor_merged['Cluster Labels'] == 1, tor_merged.columns[[2] + list(range(6, tor_merged.shape[1]))]]
```

	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	11th Most Common Venue	12th Most Common Venue
0	St. James Town, Cabbagetown	Restaurant	Italian Restaurant	Chinese Restaurant	Indian Restaurant	Thai Restaurant	Taiwanese Restaurant	Sri Lankan Restaurant	Japanese Restaurant	Caribbean Restaurant	Doner Restaurant	Fast Food Restaurant	Falafel Restaurant
3	Garden District, Ryerson	Italian Restaurant	Japanese Restaurant	Middle Eastern Restaurant	Fast Food Restaurant	Ramen Restaurant	Vietnamese Restaurant	New American Restaurant	Ethiopian Restaurant	Mexican Restaurant	Chinese Restaurant	Modern European Restaurant	Restaurant
6	Central Bay Street	Italian Restaurant	Japanese Restaurant	Indian Restaurant	Chinese Restaurant	French Restaurant	Vegetarian / Vegan Restaurant	Korean Restaurant	Middle Eastern Restaurant	Falafel Restaurant	Modern European Restaurant	Thai Restaurant	Ramen Restaurant
8	Harbourfront East, Union Station, Toronto Islands	Italian Restaurant	Restaurant	Indian Restaurant	Vegetarian / Vegan Restaurant	Sushi Restaurant	Seafood Restaurant	Japanese Restaurant	New American Restaurant	Chinese Restaurant	Doner Restaurant	Fast Food Restaurant	Falafel Restaurant
11	University of Toronto, Harbord	Restaurant	Italian Restaurant	Japanese Restaurant	Comfort Food Restaurant	Chinese Restaurant	French Restaurant	Sushi Restaurant	Brazilian Restaurant	Eastern European Restaurant	Filipino Restaurant	American Restaurant	Asian Restaurant
13	Stn A PO Boxes	Seafood Restaurant	Restaurant	Italian Restaurant	Japanese Restaurant	French Restaurant	American Restaurant	Thai Restaurant	Sushi Restaurant	Vegetarian / Vegan Restaurant	Molecular Gastronomy Restaurant	Comfort Food Restaurant	Eastern European Restaurant

CLUSTER 3

```
# Cluster 3
tor_merged.loc[tor_merged['Cluster Labels'] == 2, tor_merged.columns[[2] + list(range(6, tor_merged.shape[1]))]]
```

	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	11th Most Common Venue	12th Most Common Venue	
15	Christie	Italian Restaurant	Restaurant	Vietnamese Restaurant	Doner Restaurant	French Restaurant	Filipino Restaurant	Fast Food Restaurant	Falafel Restaurant	Ethiopian Restaurant	Eastern European Restaurant	Dumpling Restaurant	Dim Sum Restaurant	1

CLUSTER 4

```
# Cluster 4
tor_merged.loc[tor_merged['Cluster Labels'] == 3, tor_merged.columns[[2] + list(range(6, tor_merged.shape[1]))]]
```

	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	11th Most Common Venue	12th Most Common Venue	
16	Queen's Park, Ontario Provincial Government	Sushi Restaurant	Italian Restaurant	Mexican Restaurant	Vietnamese Restaurant	Doner Restaurant	French Restaurant	Filipino Restaurant	Fast Food Restaurant	Falafel Restaurant	Ethiopian Restaurant	Eastern European Restaurant	Dumpling Restaurant	1

CLUSTER 5

```
# Cluster 5
tor_merged.loc[tor_merged['Cluster Labels'] == 4, tor_merged.columns[[2] + list(range(6, tor_merged.shape[1]))]]
```

	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	11th Most Common Venue	12 C
1	Church and Wellesley	Sushi Restaurant	Japanese Restaurant	Restaurant	Mediterranean Restaurant	Indian Restaurant	Mexican Restaurant	American Restaurant	Caribbean Restaurant	Ethiopian Restaurant	Fast Food Restaurant	Afghan Restaurant	Re
4	St. James Town	American Restaurant	Italian Restaurant	Moroccan Restaurant	Restaurant	Japanese Restaurant	Vegetarian / Vegan Restaurant	Comfort Food Restaurant	Middle Eastern Restaurant	New American Restaurant	German Restaurant	Seafood Restaurant	Re
5	Berczy Park	Seafood Restaurant	Restaurant	Thai Restaurant	Greek Restaurant	Sushi Restaurant	Vegetarian / Vegan Restaurant	Japanese Restaurant	French Restaurant	Comfort Food Restaurant	Eastern European Restaurant	Vietnamese Restaurant	D Re
7	Richmond, Adelaide, King	Restaurant	Thai Restaurant	American Restaurant	Sushi Restaurant	Modern European Restaurant	Gluten-free Restaurant	Vegetarian / Vegan Restaurant	Latin American Restaurant	Mediterranean Restaurant	Colombian Restaurant	Fast Food Restaurant	A Re
9	Toronto Dominion Centre, Design Exchange	Restaurant	American Restaurant	Japanese Restaurant	Italian Restaurant	Seafood Restaurant	Asian Restaurant	Sushi Restaurant	Chinese Restaurant	Fast Food Restaurant	Vegetarian / Vegan Restaurant	French Restaurant	A Re
10	Commerce Court, Victoria Hotel	Restaurant	American Restaurant	Seafood Restaurant	Italian Restaurant	Japanese Restaurant	Thai Restaurant	Vegetarian / Vegan Restaurant	French Restaurant	Latin American Restaurant	Asian Restaurant	Gluten-free Restaurant	A Re
12	Kensington Market, Chinatown, Grange Park	Vietnamese Restaurant	Vegetarian / Vegan Restaurant	Mexican Restaurant	Comfort Food Restaurant	Japanese Restaurant	Filipino Restaurant	Dumpling Restaurant	Dim Sum Restaurant	Doner Restaurant	Caribbean Restaurant	Belgian Restaurant	Co Re
14	First Canadian Place, Underground city	Restaurant	Japanese Restaurant	American Restaurant	Asian Restaurant	Seafood Restaurant	Thai Restaurant	Sushi Restaurant	Colombian Restaurant	Fast Food Restaurant	Mediterranean Restaurant	Vegetarian / Vegan Restaurant	Re

Conclusion

Hence by using publically available data we analyzed the best areas for a Mexican Restaurant in Toronto Downtown area.

The focus of this project waas to explore the Downtown Toronto neighborhoods that will be extracted and analyzed accordingly. By using the Foursquare API we obtained the geographical location data for Downtown Toronto, and explored the restaurant venues in the neighbourhoods.

Now that the places for Venues for Mexican eatery in Downtown is mapped it will be utilized by my client for his Business expansion.

.