

COURSERA CAPSTONE PROJECT- THE BATTLE OF NEIGHBORHOODS

Himadri Bikash Das



Topics:

Introduction to problem statement

Process Flow

Analysis of the data sets and evaluation

Observation

Conclusion



Introduction to problem statement:

- Explore neighborhoods of Toronto for a best potential site to open a new restaurant.
- Evaluate the places depending upon many crucial factors such as population, distance from city center and other restaurants to minimize the competition still maintain the reach for maximum crowd to access.
- Know the most popular restaurants around and the categories so to introduce something unique and gain maximum profit.

Process Flow:

- Using the following data sources:
 - 1.Link: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
 - 2.Link: <https://foursquare.com/explore?mode=url&ne=44.418088%2C78.362732&q=Restaurant&sw=42.742978%2C-80.554504>
- Extract the pin code of all the places from the list in wiki page through web scrapping. Clean up the data and extract the neighborhood places of Toronto. Also get the latitude and longitude of the places and form a dataset.
- Use Foursquare API to get the details of the venues in and around the pin codes as per the dataset prepared earlier.
- With the help of K-Means clustering group the data for further evaluation showing the most popular venues and the existing type of restaurants.

Extracted table with the details:

- Here we can see that we have the data set ready having the Postal codes for all the neighborhoods of Toronto.
- Also we have added the Latitude and Longitude for each of these locations which will help in getting the venue details in and around these locations.

| | Postalcode | Borough | Neighborhood | Latitude | Longitude |
|----|------------|-----------------------|---|-----------|------------|
| 0 | M5B | Downtown Toronto | Garden District, Ryerson | NaN | NaN |
| 1 | M5C | Downtown Toronto | St. James Town | 43.669403 | -79.372704 |
| 2 | M4E | East Toronto | The Beaches | 43.671024 | -79.296712 |
| 3 | M5E | Downtown Toronto | Berczy Park | 43.647984 | -79.375396 |
| 4 | M5G | Downtown Toronto | Central Bay Street | 43.643430 | -79.379059 |
| 5 | M6G | Downtown Toronto | Christie | 43.664111 | -79.418405 |
| 6 | M6H | West Toronto | Dufferin, Dovercourt Village | 43.660202 | -79.435719 |
| 7 | M4J | East YorkEast Toronto | The Danforth East | 43.688171 | -79.301816 |
| 8 | M6J | West Toronto | Little Portugal, Trinity | NaN | NaN |
| 9 | M5K | Downtown Toronto | Toronto Dominion Centre, Design Exchange | NaN | NaN |
| 10 | M5L | Downtown Toronto | Commerce Court, Victoria Hotel | NaN | NaN |
| 11 | M4M | East Toronto | Studio District | 43.643059 | -79.402600 |
| 12 | M4N | Central Toronto | Lawrence Park | 43.729199 | -79.403252 |
| 13 | M5N | Central Toronto | Roselawn | 43.705998 | -79.421235 |
| 14 | M4P | Central Toronto | Davisville North | 43.697936 | -79.397291 |
| 15 | M5P | Central Toronto | Forest Hill North & West | 43.701028 | -79.425600 |
| 16 | M4R | Central Toronto | North Toronto West | 43.646547 | -79.419526 |
| 17 | M6R | West Toronto | Parkdale, Roncesvalles | 43.641281 | -79.451599 |
| 18 | M4S | Central Toronto | Davisville | 43.697936 | -79.397291 |
| 19 | M6S | West Toronto | Runnymede, Swansea | 43.651778 | -79.475923 |
| 20 | M4T | Central Toronto | Moore Park, Summerhill East | 43.686053 | -79.376519 |
| 21 | M5V | Downtown Toronto | CN Tower, King and Spadina, Railway Lands, Har... | NaN | NaN |
| 22 | M4W | Downtown Toronto | Rosedale | 43.678356 | -79.380746 |

Using Foursquare API getting the venue list:

| | Neighborhood | Neighborhood Latitude | Neighborhood Longitude | Venue | Venue Latitude | Venue Longitude | Venue Category |
|---|----------------|-----------------------|------------------------|------------------------|----------------|-----------------|----------------------|
| 0 | St. James Town | 43.669403 | -79.372704 | Mr. Jerk | 43.667328 | -79.373389 | Caribbean Restaurant |
| 1 | St. James Town | 43.669403 | -79.372704 | Cranberries | 43.667843 | -79.369407 | Diner |
| 2 | St. James Town | 43.669403 | -79.372704 | Murgatroid | 43.667381 | -79.369311 | Restaurant |
| 3 | St. James Town | 43.669403 | -79.372704 | Absolute Bakery & Café | 43.667469 | -79.369277 | Bakery |
| 4 | St. James Town | 43.669403 | -79.372704 | F'Amelia | 43.667536 | -79.368613 | Italian Restaurant |

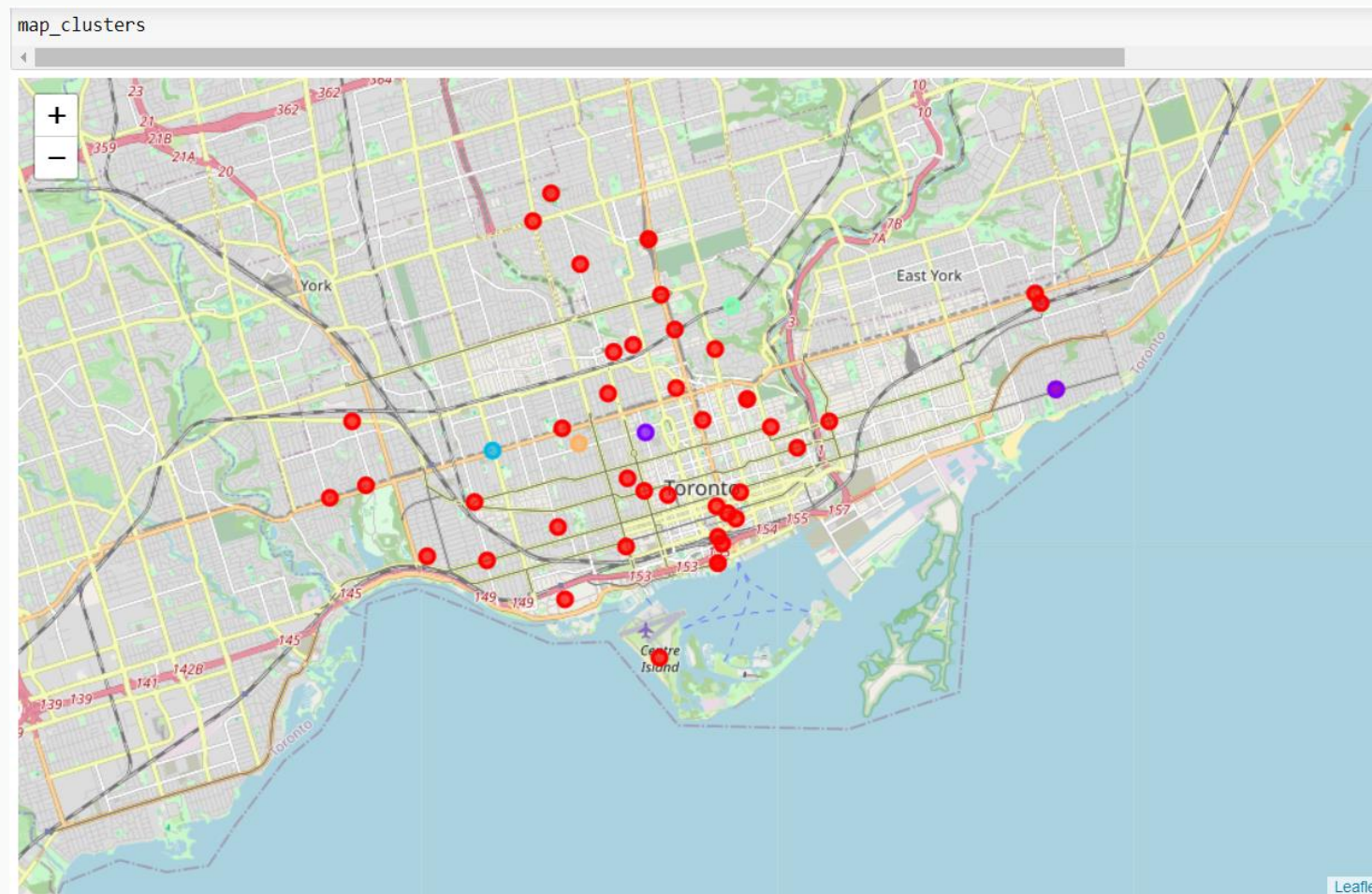
- Here we can see that from Foursquare API the venues for the corresponding neighborhood has been extracted with their Latitude, Longitude as well as the venue category.
- We can further group the venues category wise and evaluate the upcoming clusters for detailed analysis

Extracting the most popular venues for each neighborhood:

- Now we have segregated the most popular venues for each neighborhood.
- This analysis will help us evaluating the most popular cafes and restaurants and also their types.
- This will essentially help with decision making as in what kind of restaurant we are planning to open.

| | Neighborhood | 1st Popular Venues | 2nd Popular Venues | 3rd Popular Venues | 4th Popular Venues | 5th Popular Venues | 6th Popular Venues | 7th Popular Venues | 8th Popular Venues | 9th Popular Venues | 10th Popular Venues |
|---|------------------|---------------------|--------------------|--------------------------|-------------------------|---------------------|----------------------|----------------------|---------------------|--------------------|-------------------------------|
| 0 | Adelaide | Coffee Shop | Café | Hotel | Gym | Restaurant | Japanese Restaurant | Italian Restaurant | American Restaurant | Clothing Store | Gastropub |
| 1 | Cabbagetown | Restaurant | Café | Coffee Shop | Japanese Restaurant | Italian Restaurant | Indian Restaurant | Chinese Restaurant | Pub | Pizza Place | Diner |
| 2 | Chinatown | Café | Bar | Dessert Shop | Comfort Food Restaurant | Clothing Store | Ice Cream Shop | Coffee Shop | Mexican Restaurant | Park | Vegetarian / Vegan Restaurant |
| 3 | Deer Park | Coffee Shop | Italian Restaurant | Grocery Store | Sushi Restaurant | Thai Restaurant | Pizza Place | Gym | Bank | Café | Bagel Shop |
| 4 | Exhibition Place | Arts & Crafts Store | Soccer Stadium | Poutine Place | Coffee Shop | Athletics & Sports | Music Venue | Comedy Club | Plaza | Restaurant | Flea Market |
| 5 | Forest Hill SE | Playground | Park | Mediterranean Restaurant | Bank | Filipino Restaurant | Fast Food Restaurant | Farmers Market | Falafel Restaurant | Fish & Chips Shop | Dog Run |
| 6 | Grange Park | Coffee Shop | Café | Japanese Restaurant | Sandwich Place | Clothing Store | Arts & Crafts Store | Art Gallery | Record Shop | Sushi Restaurant | French Restaurant |
| 7 | Harbord | Korean Restaurant | Café | Grocery Store | Coffee Shop | Pizza Place | Karaoke Bar | Fast Food Restaurant | Sandwich Place | Bar | Mexican Restaurant |
| 8 | Harbourfront | Coffee Shop | Café | Restaurant | Hotel | Pizza Place | Italian Restaurant | Park | Bank | Music Venue | Chinese Restaurant |

Resulting cluster on map:



Observation:

Evaluating the clusters formed we can conclude the following:

- a. Cluster 1 : This cluster includes most of the neighborhoods. Out of the categories we can see that the coffee shop been the most popular venue among these places followed by Korean and Italian restaurants.

So we can say clearly that coming up with a café as an option for new business will attract more customer since these places are popular with coffee but also have to face a considerable amount of competition.

- b. Cluster 2 : This comprises of less neighborhood and café and pizza amounts to the most popular venue.

Observation:

Evaluating the clusters formed we can conclude the following:

- c. Cluster 3 : This is a very localized one with just one neighborhood and bar being the most popular venue.
- d. Cluster 4 & 5: Both of these clusters consist of just one neighborhood again having park and Korean restaurant respectively as the popular destination.

Conclusion:

Finally after evaluating different clusters we can conclude that Cluster 4 does not have any prominent restaurant nearby but since the most popular venue is park we can expect lots of people visiting the place.

Hence the idea of opening up a restaurant will definitely help people. And moreover as with the Other clusters we can see that café and Italian food is a great choice for the people of many Neighborhoods of Toronto so may be the idea of a café or Italian restaurant will surely click.

THANK YOU

