Finding the best location for a Mexican restaurant in Toronto Canada using data science analysis

1. **Introduction**
    a. **Background**

    We want to open a Mexican restaurant in Toronto Canada, its one of the most popular food around the world and maybe in a certain area of Toronto there not to many restaurant of this type, so the idea is to find the best place to put a restaurant of this type.

    b. **Business Problem**

    The aim of the project is to find the best location for a Mexican food restaurant using various tools and techniques such as data science methodology, machine learning algorithms such as clustering and tools like Foursquare to gather the necessarily data for the analysis

    With this project we are answering the question

    Which will be the best location to open a Mexican food restaurant in Toronto Canada?

2. **Data**
    a. The data that we will use to answer the question is:
        i. List of postal codes of Canada: M
        ii. CSV with the coordinate of every Borough in Toronto Canada
        iii. Information gathered from foursquare of the categories for every Venue.
3. **Methodology**
    a. First we do a web scrapping to get the data from wikipedia, we will get the Lis of postal codes of Canada in Toronto area, for this part we are using the library beautiful soup for the process to be more easy, then we do a cleaning of the data drooping the values in borough that appears as ***"Not assigned"*** , after this process we need to load the csv file containing the coordinates data for each PostalCode,

we will need this data to create the map and the clusters and also to gather the data from foursquare api.
After this we make a join between the two data frames to get the coordinates and also the the boroughs, finally we create a map to visualize that everything is working correctly

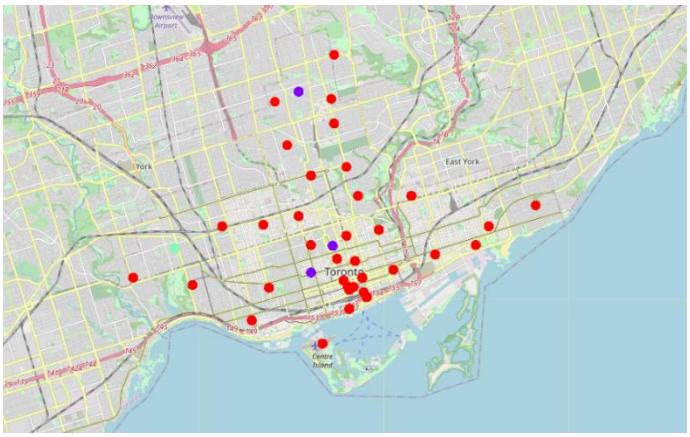
Then we will use foursquare to get the top 100 venues in a radius of 500 meters from every one of the Neighborhoods in Toronto
Once done this lets clean the data and process it, in this case we are focusing on Mexican restaurants, so let's filter the data to see how many Mexican restaurant has every neighborhood, for this part lets group the neighborhoods and get just the average of Mexican restaurants for every neighborhood.
Once done this we will create 3 clusters, one for neighboors that has none or little mexican restaurants, another one with has medium mexican restaurants and finally one we it has a lot of mexican restaurant, we are doing this with the aim of analyzing where is better to put the new Mexican Restaurant

4. **Results**

Lets analyze the map that we got

	Neighborhoods	Mexican Restaurant
6	Church and Wellesley	0.012346
13	Garden District, Ryerson	0.010000
15	High Park, The Junction South	0.086957
17	Kensington Market, Chinatown, Grange Park	0.049180
21	North Toronto West, Lawrence Park	0.052632
23	Queen's Park, Ontario Provincial Government	0.030303

As we can see in the image this Neighborhoods are the ones that has the most mexican restaurants, so since this point we can say that a new Mexican Restaurant shouldn't be in this neighborhoods



In the map we can see the different clusters

- The green cluster is the cluster number 2, here we have a lot of Mexican restaurants
- The red cluster is cluster number 0, here we don't have many Mexican restaurants
- The purple cluster is the cluster number 1, in this cluster we have some Mexican restaurants

5. Discussion

This project just considered the amount of similar restaurants that can be found in a certain area, for future research we should include some other indicators such as amount of people living there, maybe the gross income for the neighborhoods, the security of the areas, the diversity of the neighborhood etc.

5. **Conclusions**

As we can see in the image the red cluster refers to the neighbors that doesn't have a lot of Mexican restaurant, for our purpose we should select one of this neighbors to establish our Mexican restaurant

	Neighborhood	Mexican Restaurant	Cluster Labels	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Berczy Park	0.00	0	43.644771	-79.373306	LCBO	43.642944	-79.372440	Liquor Store
28	Runnymede, Swansea	0.00	0	43.651571	-79.484450	A Dark Horse	43.649533	-79.483056	Bar
28	Runnymede, Swansea	0.00	0	43.651571	-79.484450	Tim Hortons	43.648526	-79.485066	Coffee Shop
28	Runnymede, Swansea	0.00	0	43.651571	-79.484450	Heart	43.650310	-79.480125	Dessert Shop
28	Runnymede, Swansea	0.00	0	43.651571	-79.484450	Java Joe's	43.650424	-79.479755	Café
...	...	...	...	...	...	...	...	...	...
13	Garden District, Ryerson	0.01	0	43.657162	-79.378937	Taiyaki NYC - Toronto	43.655580	-79.384462	Ice Cream Shop
13	Garden District, Ryerson	0.01	0	43.657162	-79.378937	Urban Outfitters	43.654411	-79.380055	Clothing Store
13	Garden District, Ryerson	0.01	0	43.657162	-79.378937	Vans	43.654826	-79.380241	Shoe Store
13	Garden District, Ryerson	0.01	0	43.657162	-79.378937	Ali Basha Café	43.656690	-79.375459	Hookah Bar
4	Central Bay Street	0.00	0	43.657952	-79.387383	Uncle Tetsu's Cheesecake (Uncle Tetsu's Japane...	43.656063	-79.383695	Dessert Shop

So to answer our question, we should select one of the following neighborhoods to establish our Mexican Restaurant.

'Berczy Park', 'Runnymede, Swansea', 'St. James Town', 'Richmond, Adelaide, King', 'Roselawn','Rosedale', 'St. James own, Cabbagetown', 'Little Portugal, Trinity', 'Moore Park, Summerhill East', 'Regent Park, Harbourfront', 'Parkdale, Roncesvalles', 'University of Toronto, Harbord', 'India Bazaar, The Beaches West', 'Lawrence Park', 'Toronto Dominion Centre, Design Exchange', 'The Danforth West, Riverdale', 'The Beaches', 'The Annex, North Midtown, Yorkville', 'Stn A PO Boxes', 'Summerhill West, Rathnelly, South Hill, Forest Hill SE, Deer Park', 'Studio District',

'First Canadian Place, Underground city', 'Davisville', 'Commerce Court, Victoria Hotel', 'Davisville North', 'Dufferin, Dovercourt Village', 'Central Bay Street', 'Church and Wellesley', 'Christie', 'Harbourfront East, Union Station, Toronto Islands', 'CN Tower, King and Spadina, Railway Lands, Harbourfront West, Bathurst Quay, South Niagara, Island airport', 'Business reply mail Processing Centre, South Central Letter Processing Plant Toronto',

'Brockton, Parkdale Village, Exhibition Place', 'Garden District, Ryerson', 'Forest Hill North & West, Forest Hill Road Park'