

Opening a new Bermese Restaurant in Toronto

Burmese Restaurant in Toronto

- There may not be enough Burmese restaurants in Toronto and it might present a great opportunity for this entrepreneur who is based in Canada.
- As Burmese food is very similar to other Asian cuisines, as Canada has a lot of migrant from Asia so finding a prime location to open a restaurant with such cuisine will be beneficial for any entrepreneur.
- To find the most suitable location for the entrepreneur to open a new Burmese restaurant in Toronto, Canada. By using data science methods and machine-learning methods such as clustering, this project aims to provide solutions.

Data acquisition

To solve this problem, We will need below data

- List of neighborhoods in Toronto, Canada.
- Latitude and Longitude of these neighborhoods.
- Venue data related to Asian restaurants.
- List of neighborhoods in Toronto

https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

Extracting the Data:

- The scrapping of Toronto neighborhoods via Wikipedia
- Getting Latitude and Longitude data of these neighborhoods via Geocoder package
- Using Foursquare API to get venue data related to these neighborhoods

Methodology

- Extracting the list of neighborhoods from Wikipedia page (“https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M”)
- To get their coordinates to utilize Foursquare to pull the list of venues near these neighborhoods. To get the coordinates, I tried using Geocoder package but it was not working so I used the CSV file provided by the IBM team to match the coordinates of Toronto neighborhoods.
- From Foursquare, I can pull the names, categories, latitude, and longitude of the venues. With this data, I can also check how many unique categories that I can get from these venues.
- I analyze each neighborhood by grouping the rows by neighborhood and taking the mean on the frequency of occurrence of each venue category

Methodology

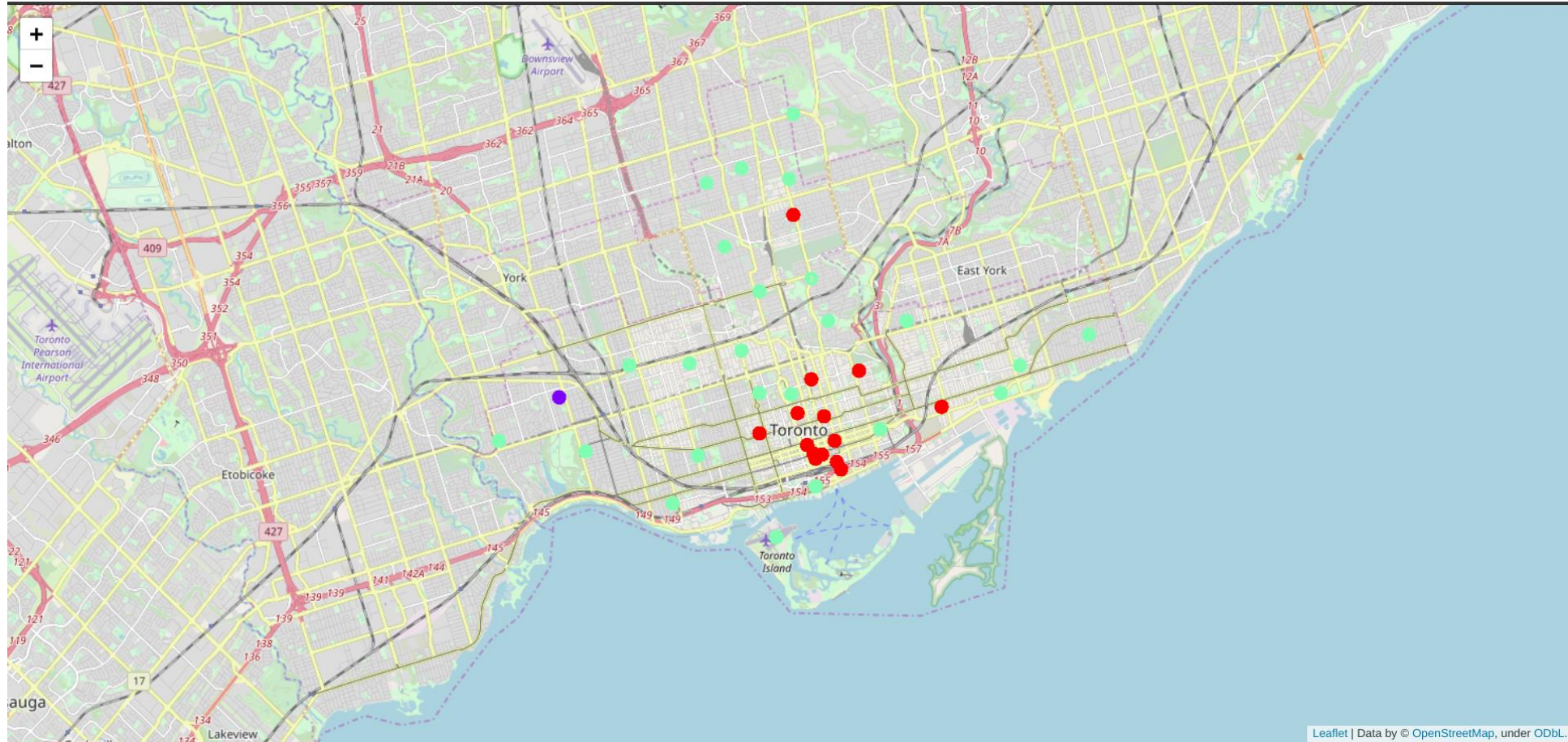
- when I ran the model, I was looking for “Asian restaurants” but there are very few results (maybe due to Foursquare categorization) so I looked for the restaurants closest to Burmese cuisine taste.
- I performed the clustering method by using k-means clustering. K-means clustering algorithm identifies k number of centroids, and then allocates every data point to the nearest cluster while keeping the centroids as small as possible.
- I have clustered the neighborhoods in Toronto into 3 clusters based on their frequency of occurrence for “Thai food”. Based on the results (the concentration of clusters), I will be able to recommend the ideal location to open the restaurant.

Results

The results from k-means clustering show that we can categorize Toronto neighborhoods into 3 clusters based on how many Thai restaurants are in each neighborhood

- Cluster 0: Neighborhoods with little or no Thai restaurants
- Cluster 1: Neighborhoods with no Thai restaurants
- Cluster 2: Neighborhoods with a high number of Thai restaurants

Cont..



Limitations and future work

- Consideration of one factor: the occurrence/existence of Thai restaurants in each neighborhood. Many factors can be taken into consideration such as population density, the income of residents, rent that could influence the decision to open a new restaurant.
- Future research can take into consideration these factors. Also, I am relying on the existence of Thai restaurants only for this project but future research can take into consideration other variables such as the existence of Asian restaurants

Conclusion

We have gone through the process of identifying the business problem, specifying the data required, extracting and preparing the data, performing the machine learning by utilizing k-means clustering and providing recommendations to the stakeholder.