

Recommendation System - Starting a New Restaurant Business

Abhinay Dornipati

Table of Contents

- Introduction/Business Problem
- Data
- Methodology
- Analysis
- Results and Discussion
- Conclusion

Introduction/Business Problem

Problem Background:

- Recommendation model to recommend an optimal location for a new restaurant business in the city of Toronto.
- For instance, to recommend an optimal location, ABC Company Ltd has appointed me to lead the Data Science Team.
- The goal is to locate and recommend which neighborhood of Toronto would be the best choice to start a restaurant serving a specific type of cuisine.

Problem Description:

Factors considered to decide upon the location and the type cuisine to be served in the restaurant are;

- Toronto Population and demographics
- Competitors in that locality
- The cuisine served by the competitors
- Study about the popular venues around that locality – for example, Tourist places, Parks, Theaters etc.
- Study of the locations or neighborhoods based on the Boroughs in which they are present.

Data

For building the recommendation model, the following data and information is considered for analysis

- Scrapped Wikipedia using **BeautifulSoup**, to extract information about **10 Toronto boroughs**, also known as local authority districts. Also, the local areas or **neighborhoods** for each borough are considered for detailed analysis.
- **Foursquare API** is used to get information about the available restaurants in a given neighborhood and given borough. The API also provides information about the category of restaurants based on the cuisine each restaurant is serving.
- The **location data** or address in the form of latitude and longitude was present for each neighborhood in the form of a **.csv** file.

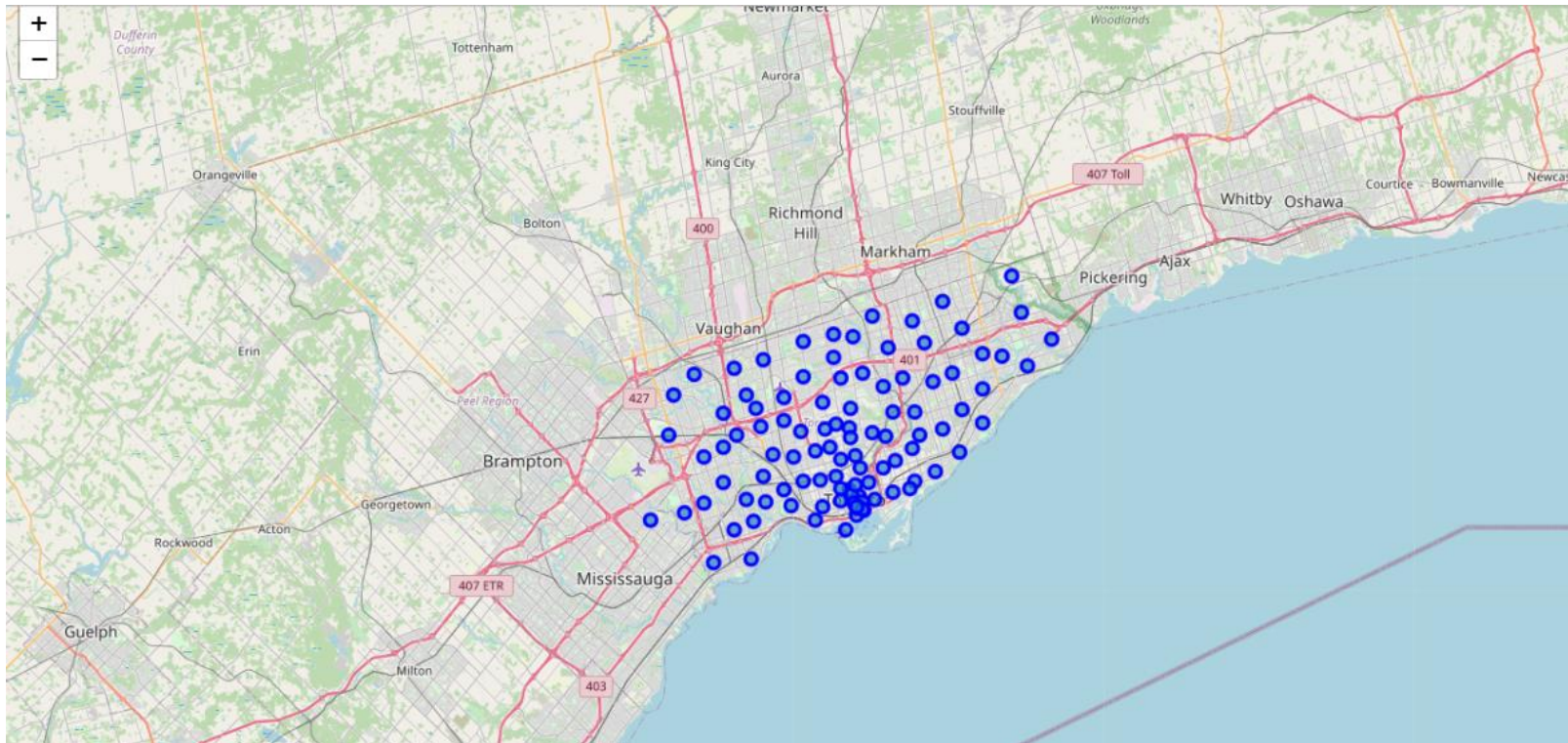
The link to the Wikipedia page is given below:

[Tononto_Boroughs_Neighborhoods](#)

Methodology

Exploratory Data Analysis:

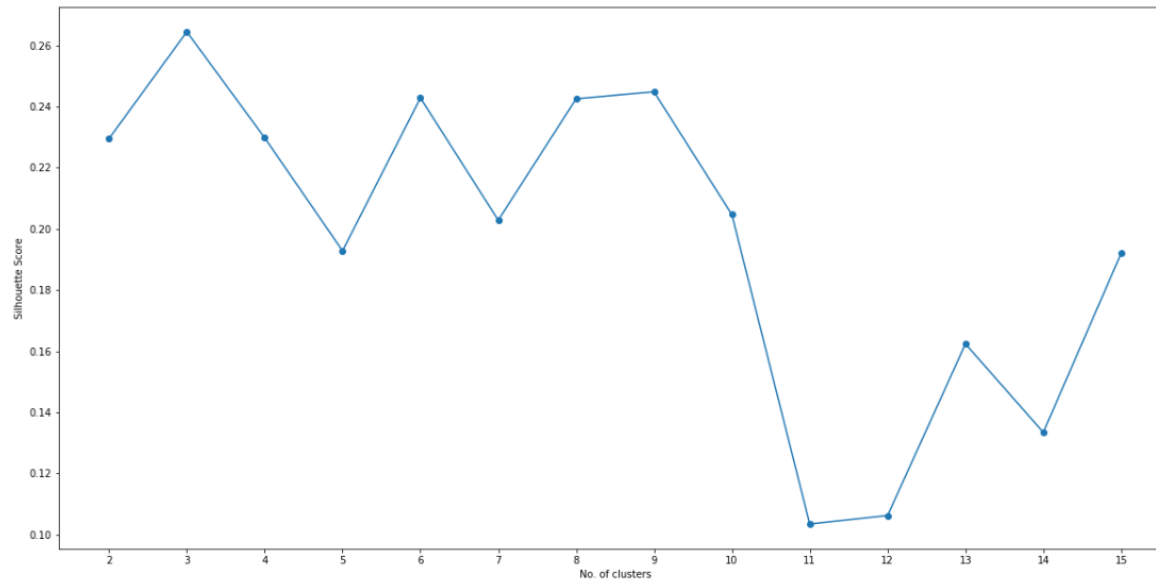
- Toronto is mainly divided into 10 boroughs also known as local authority districts with 103 neighborhoods. Using **Folium** library, map of Toronto along with its neighborhoods is created for a better visualization.



Methodology

Problem Approach and Employment of k-means Clustering:

- k-means clustering is performed on the data frame to check the pattern for each neighborhood.
- The best value of 'k', i.e, the number of clusters is determined using silhouette_score.
- A graph has been plotted between the silhouette_score and different values of k.
- Number of clusters (k) is chosen to be 3.



Methodology

Problem Approach and Employment of k-means Clustering:

- Each neighborhood in the region of Toronto is assigned to one of the clusters based on the top ten restaurants serving same cuisines in the respective neighborhood.
- The “Cluster Labels” column is added to the data frame, where 0 represents 1st cluster, 1 represents 2nd cluster and 2 represents the 3rd cluster.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category	Cluster Labels	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
5	Victoria Village	43.725882	-79.315572	Portugril	43.725819	-79.312785	Portuguese Restaurant	0	Portuguese Restaurant	Vietnamese Restaurant	Doner Restaurant	German Restaurant	French Restaurant	Filipino Restaurant	Fast Food Restaurant	Falafel Restaurant	Ethiopian Restaurant	Empanada Restaurant
13	Regent Park / Harbourfront	43.654260	-79.360636	Impact Kitchen	43.656369	-79.356980	Restaurant	0	Restaurant	Asian Restaurant	French Restaurant	Mexican Restaurant	Vietnamese Restaurant	Dumpling Restaurant	German Restaurant	Filipino Restaurant	Fast Food Restaurant	Falafel Restaurant
29	Regent Park / Harbourfront	43.654260	-79.360636	Cluny Bistro & Boulangerie	43.650565	-79.357843	French Restaurant	0	Restaurant	Asian Restaurant	French Restaurant	Mexican Restaurant	Vietnamese Restaurant	Dumpling Restaurant	German Restaurant	Filipino Restaurant	Fast Food Restaurant	Falafel Restaurant
33	Regent Park / Harbourfront	43.654260	-79.360636	Cocina Economica	43.654959	-79.365657	Mexican Restaurant	0	Restaurant	Asian Restaurant	French Restaurant	Mexican Restaurant	Vietnamese Restaurant	Dumpling Restaurant	German Restaurant	Filipino Restaurant	Fast Food Restaurant	Falafel Restaurant
41	Regent Park / Harbourfront	43.654260	-79.360636	Izumi	43.649970	-79.360153	Asian Restaurant	0	Restaurant	Asian Restaurant	French Restaurant	Mexican Restaurant	Vietnamese Restaurant	Dumpling Restaurant	German Restaurant	Filipino Restaurant	Fast Food Restaurant	Falafel Restaurant

Analysis

Neighborhood K-Means clustering based on the mean occurrence of venue category:

- All the 3 clusters have unique pattern with regards to the top ten restaurants for a particular neighborhood. The picture below presents the number of neighborhoods assigned to each cluster.

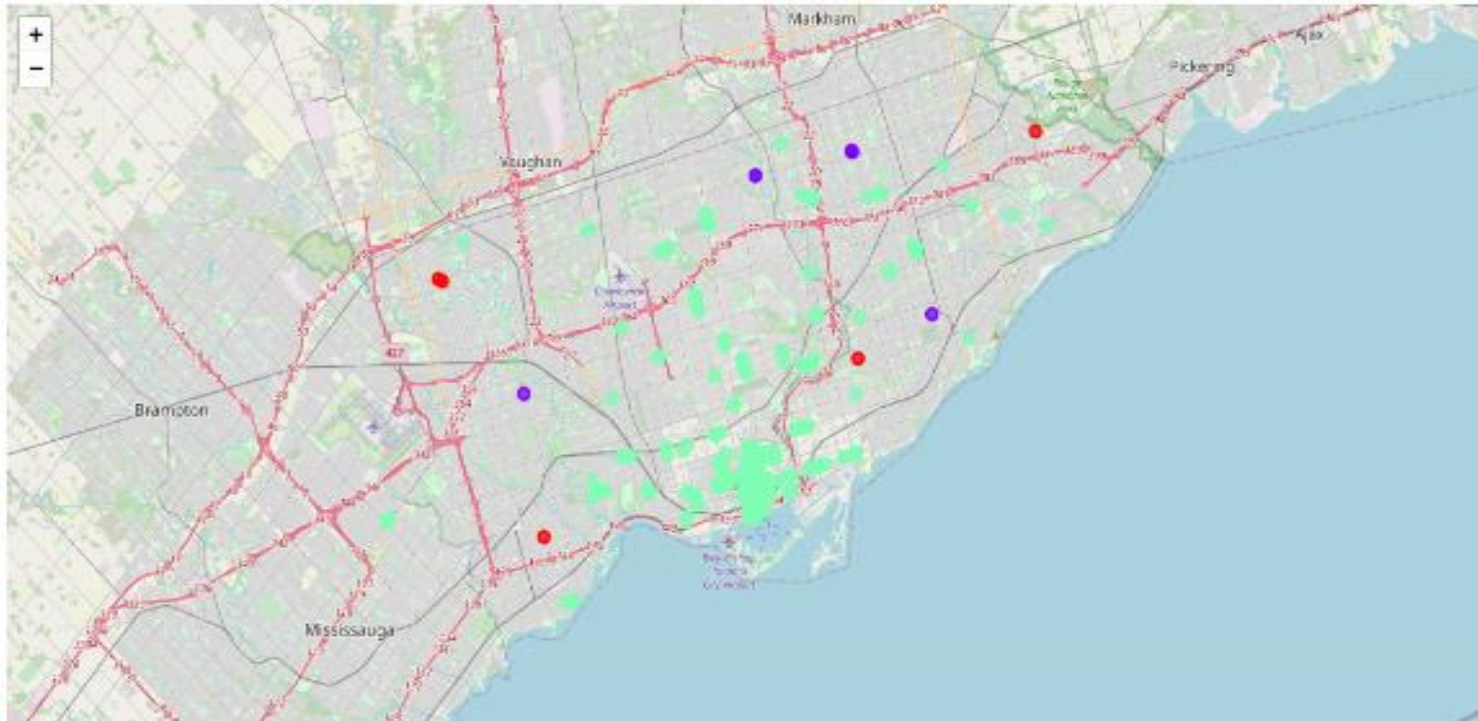
Cluster	Number of Neighborhoods
2	486
1	8
0	5

- Cluster 2 stands on the top with highest number of neighborhoods of 486. Whereas, the clusters 1 and 0 have only 8 and 5 number of neighborhoods respectively.

Analysis

Map of Toronto with clustered neighborhoods:

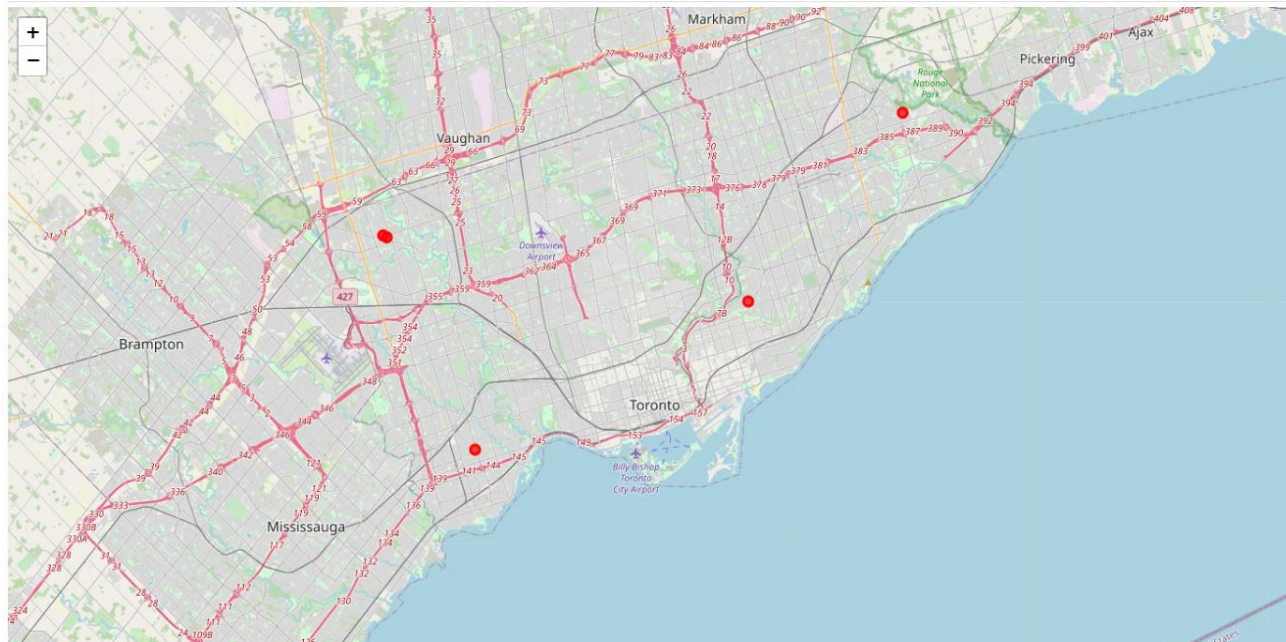
- As seen from the figure the clusters differ a lot from each other in terms of the number of neighborhoods in each of them.



Results and Discussion

Analysis: Each cluster of the neighborhoods in Toronto:

- Cluster 1

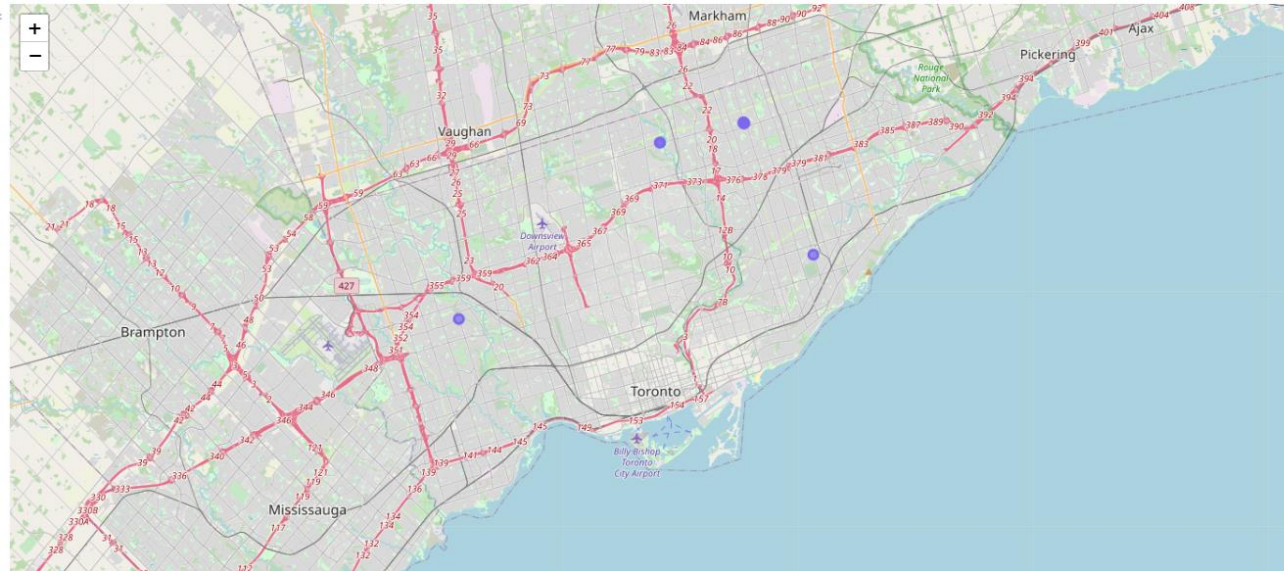


- As already mentioned there only 5 neighborhoods in this cluster.

Results and Discussion

Analysis: Each cluster of the neighborhoods in Toronto:

- Cluster 2

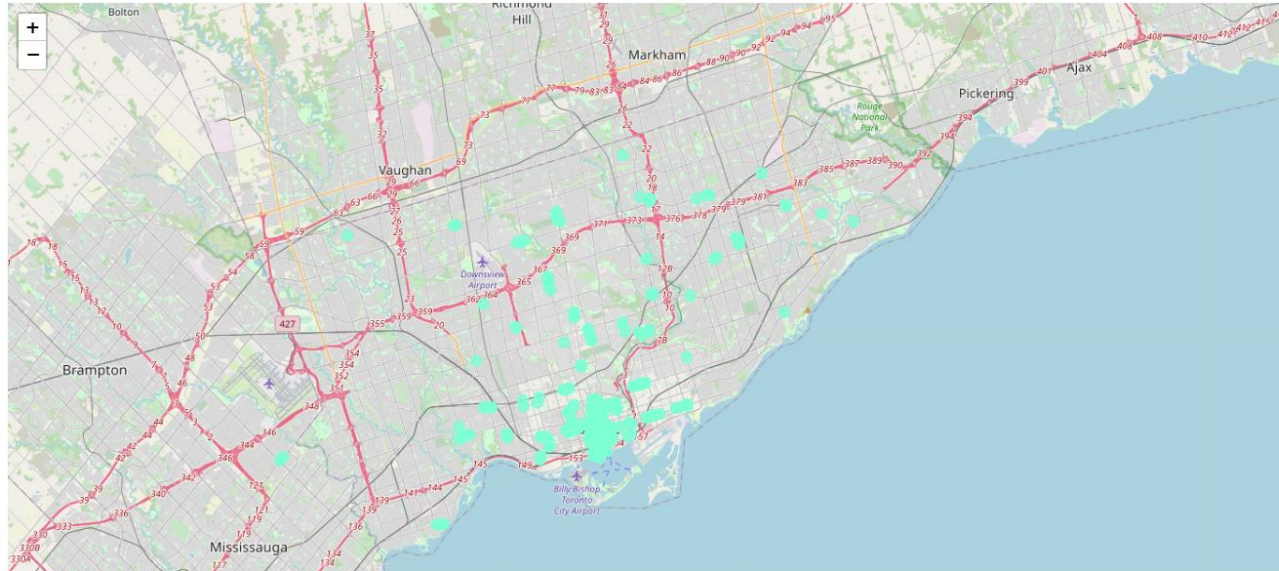


- Cluster 2 contains 8 neighborhoods in it.

Results and Discussion

Analysis: Each cluster of the neighborhoods in Toronto:

- Cluster 1



- Cluster 3 is the most populous one with 486 number of neighborhoods in it.

Results and Discussion

Comparison of the 3 Clusters:

	Cluster	Most Common Restaurant	Least Common Restaurant	2nd Least Common Restaurant
0	1	Fast Food	Eastern European Restaurant	Empanda Restaurant
1	2	Chinese Restaurant	Empanda Restaurant	Ethiopian Restaurant
2	3	Italian Restaurant	Gulden-free Restaurant	Japanese Restaurant

The results can be analyzed in two ways:

- If ABC Company Ltd wants to open a restaurant in an already decided location irrespective of the cuisine, then it's good for them to serve a cuisine which is not being served by any of the restaurants in that location. This would bring good profits to the company.
- In contrast, if the ABC Company Ltd wants to open a restaurant serving a specific cuisine irrespective of the location, then it's preferred to open the restaurant in a location where this kind of cuisine is not served or at least served by least number of restaurants.

Conclusion

Conclusions drawn from the analysis:

- This model is built on limited amount of data available. Therefore, the recommendation made might not be totally accurate. With the availability of more amount of data, this model can make better recommendations.
- It has been primarily observed that, even though Toronto has large number of restaurants, they are all present only in certain neighborhoods.
- Few neighborhoods have very less number of restaurants serving only few kinds of cuisine. Therefore, there is a lot of scope to start new restaurants in such locations, serving the cuisines not available till date.
- For example, every neighborhood can be checked for number of restaurants available and the type of cuisine being served. Accordingly, a location with lowest competition and thereby lowest risk can be chosen to start a new restaurant business.