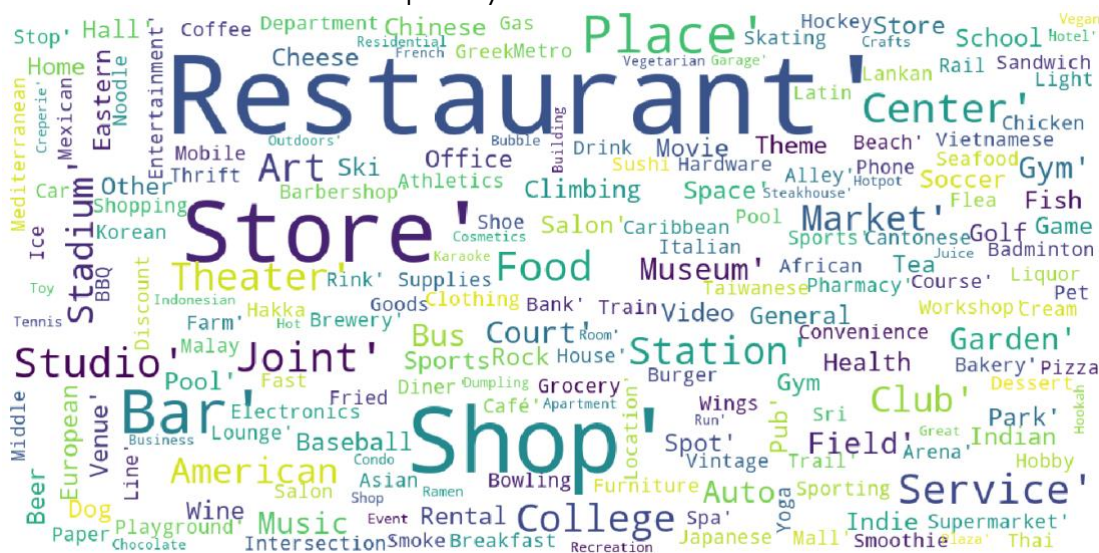


1. Introduction

1.1 Background

People love food. Ability to enjoy the taste of it has been our curse and blessing for as long as humankind is. Nowadays it is much easier to dive into different gastronomical experiences without travelling abroad. Even more, you can take pleasure of the vast majority of world cuisines not leaving your neighbourhood. And as many eating places there would be that many of them people would enjoy. Though food business always to be available and demanded.

Toronto is the densely populated area of Canada. It is the largest city in Canada with a population of over 6 million. Toronto is a place where people can try the best of each culture. In neighbourhoods like Corso Italia, Chinatown, Kensington Market, Little India, Little Italy, Koreatown and so on are plenty of restaurants to eat in.



It would be fair to notice that restaurants are rather popular venues in Toronto, so there might be a temptation to open an additional one.

1.2 Problem

The objective of this project is to determine what might be the most suitable neighbourhood of Toronto to open a restaurant using Foursquare location data and clustering of venue information. The most suitable location for a new Mexican restaurant in Toronto, Canada, will be determined as a conclusion of this project.

1.3 Interest

This project might be exciting and fruitful for those who have in mind opening a Mexican restaurant in the capital of Canada.



2. Data acquisition and cleaning

2.1 Data sources

To assess the best neighbourhood for a Mexican restaurant, we need data about the communities of Toronto and data about venues in the areas.

We will fetch data about Toronto' neighbourhoods from Wikipedia page with postal codes, boroughs, and corresponding neighbourhoods. Coordinates of areas we will get from CSV file kindly provided by this course.

Information about venues in each area will be obtained via the Foursquare API. We will take 200 places at max in a radius of 1 km from the centre of each neighbourhood.

After cleaning, it will be used to classify areas of Toronto and get the most suitable cluster for building a new Mexican restaurant.

2.2 Data cleaning

Only process the cells that have an assigned borough. Ignore cells with a borough that is Not assigned.

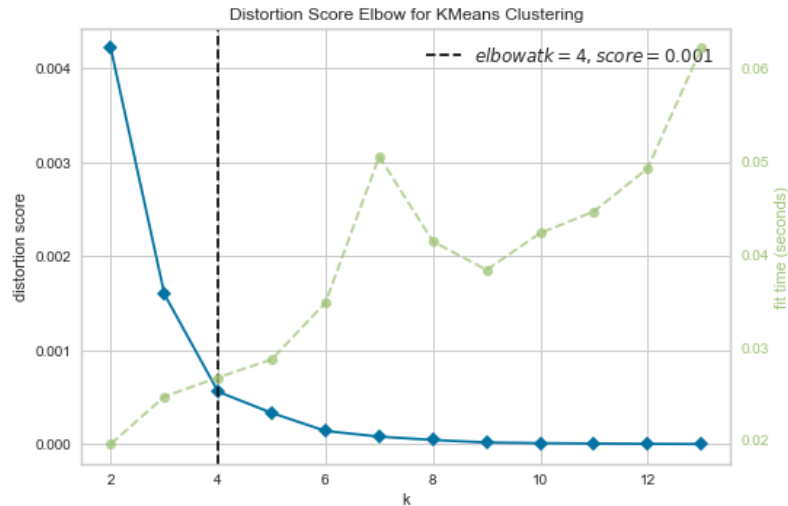
More than one neighbourhood can exist in one postal code area. For example, in the table on the Wikipedia page, you will notice that M5A is listed twice and has two neighbourhoods: Harbourfront and Regent Park. These two rows will be combined into one row with the communities separated with a comma, as shown in row 11 in the above table.

If a cell has a borough but a Not assigned neighbourhood, then the area will be the same as the borough.

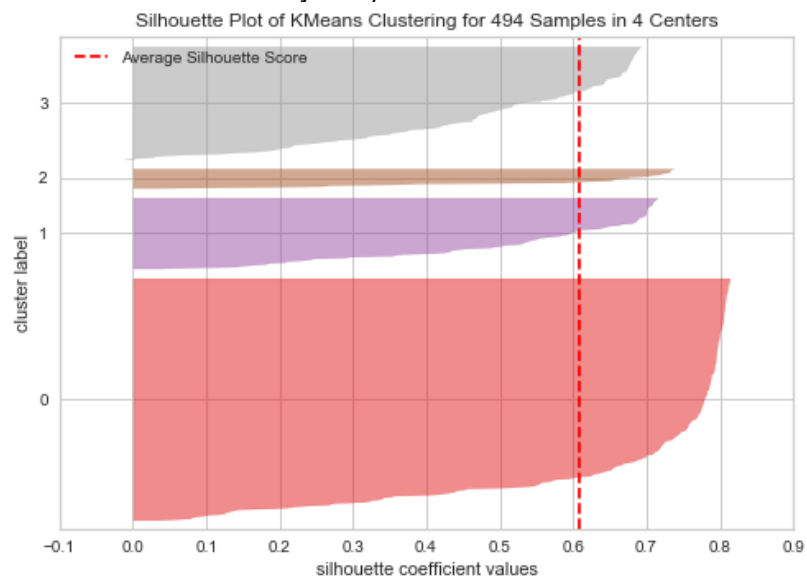


3. Methodology

We applied KMean algorithm to cluster out neighbourhoods. K was determined with the elbow method using yellowbrick. The optimum K value was 4, so areas were divided into 4 clusters.



Silhouette were used to justify it.

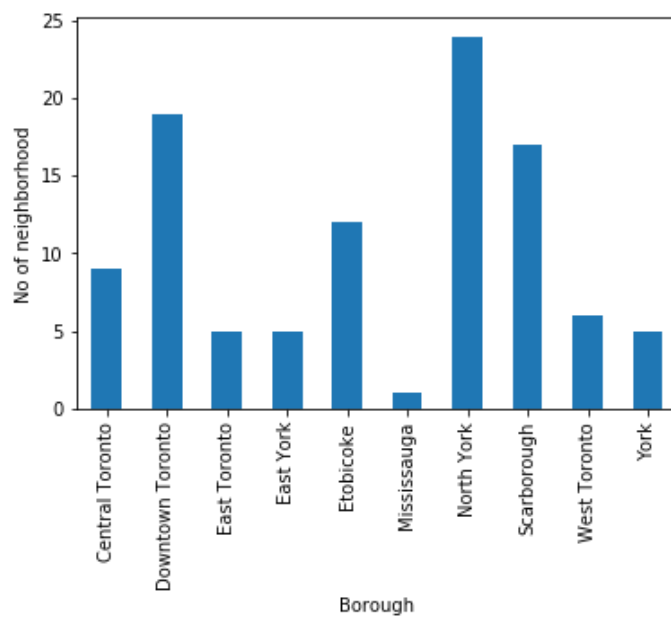
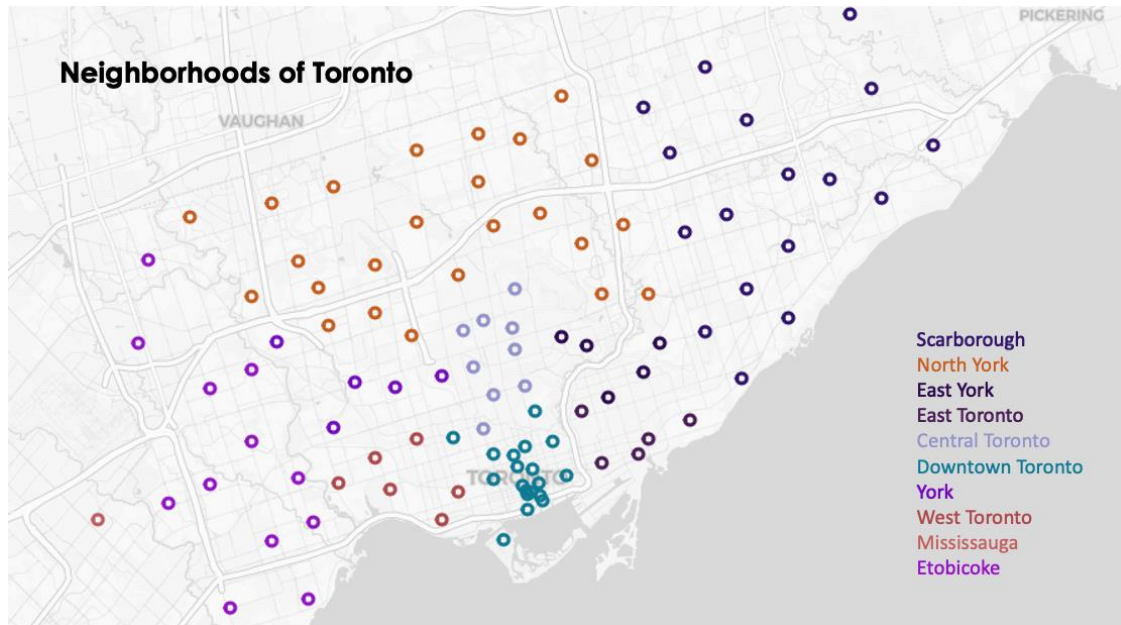


Then we define cluster with the lowest number of Mexican restaurants, aligned it with a density of areas in the cluster and made the recommendation.



4. Result

All together there where 103 neighbourhoods grouped into 10 boroughs.



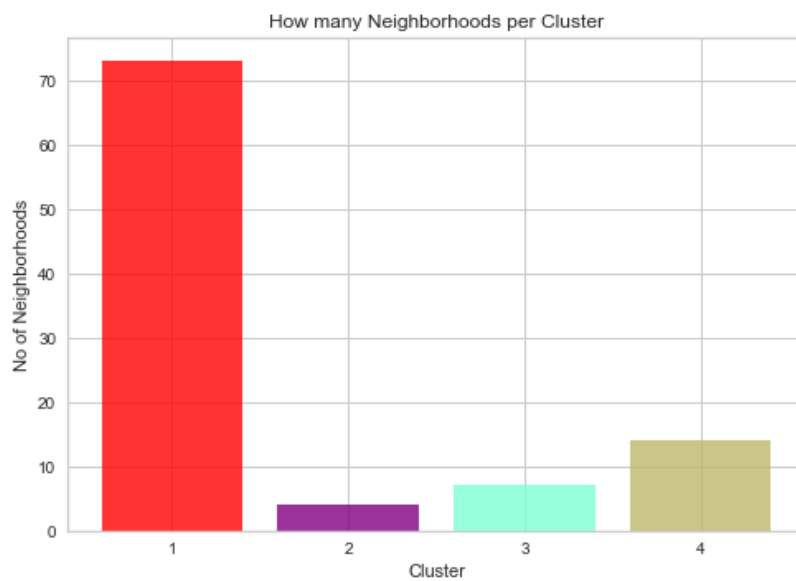
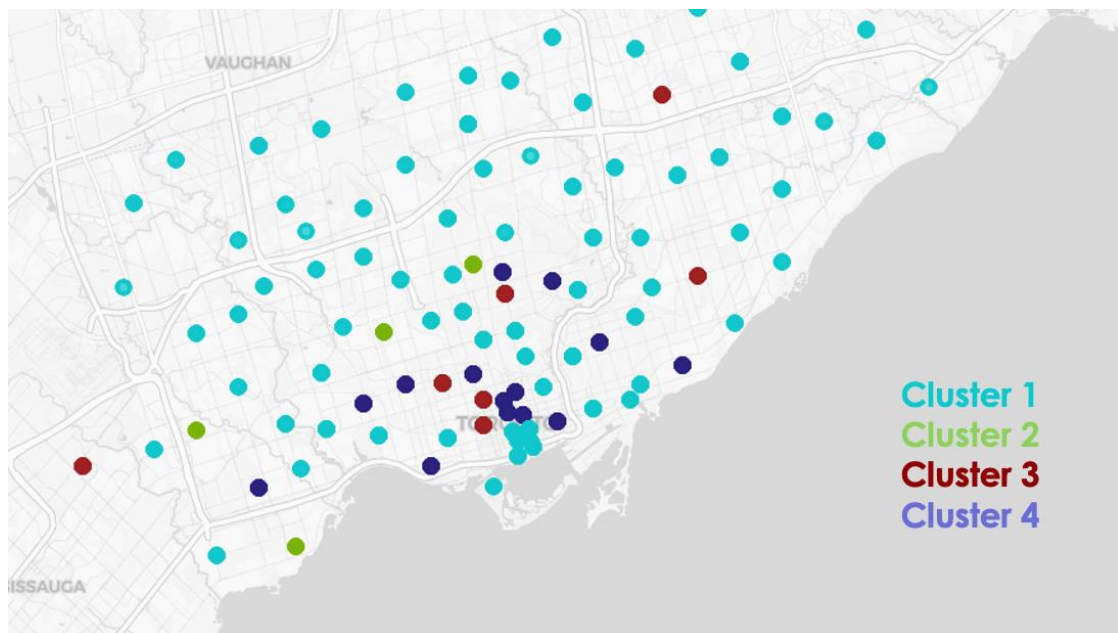
The densest were North York (24 areas), Downtown Toronto (19 areas) and Scarborough (17 areas).

After obtaining data from the Foursquare we got following.

There were 4907 venues in 329 unique categories. Over all, there were 39 Mexican restaurants.



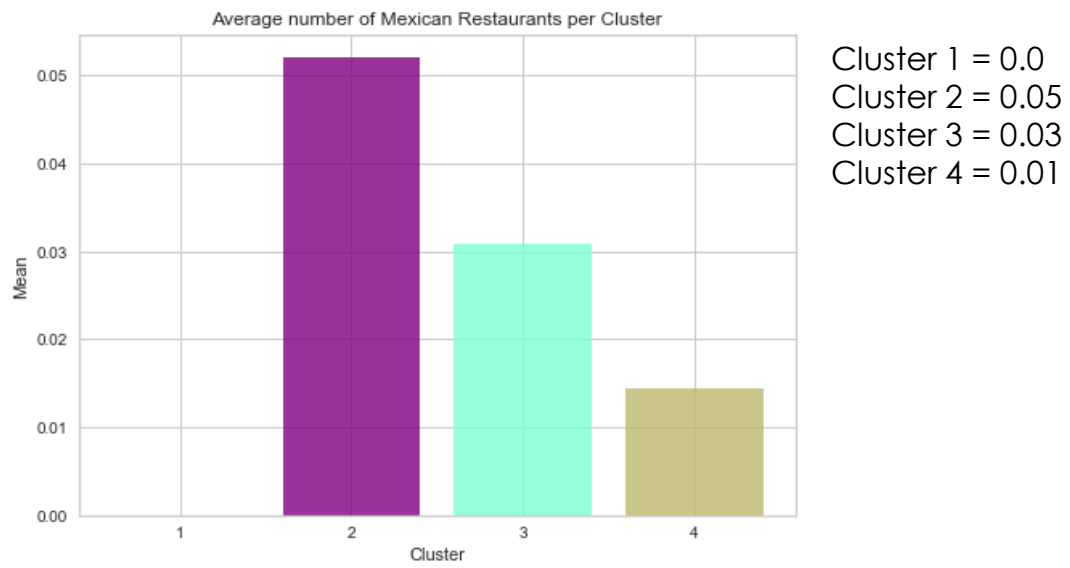
After clustering we have got 4 clusters of neighbourhoods.



Cluster 1 appeared to be the biggest one.



Then we defined rate of Mexican restaurants in each cluster.



5. Discussion

Data that we have got showed us that there are plenty of areas of Toronto that do not have Mexican restaurants at all. These areas formed cluster 1. Cluster 2, 3, and 4 had moderate number of Mexican restaurants but at same time were formed from relatively small number of neighbourhoods.

It was reasonable to assume that the most suitable area to build a Mexican restaurant would be one that had few or any of them.



6. Conclusion

The very best cluster to build a Mexican restaurant is cluster 1, because there is no Mexican restaurants at all. Furthermore, this cluster includes biggest number of neighborhoods, which are less dense, even if some new restaurant might be build there, there still will be low competition.

Any of neighborhoods would be good for a Mexican restaurant.

