

1. Introduction :

Problem background:

Toronto is Canada's Largest city and a world leader in such areas as business, finance, technology, entertainment and culture. With a recoded population of 2,731,571 in 2016, it is the most populous city in Canada and the fourth most populous city in North America. Toronto's large population of immigrants from all over the globe has also made Toronto one of the most multicultural cities in the world.

One of the best things about living in the world's most multicultural city is all the tasty and diverse food options. The diversity of the cuisine available is reflective of the social and economic diversity of Toronto. Whether you are in the mood for Italian, Mexican, Chinese or Ethiopian, you won't have a problem finding a delicious meal to satisfy your craving in Toronto.

Business Understanding/Problem Description:

In this project, we will find the best neighborhood in Toronto for opening a new Chinese restaurant.

There are some questions we need to address:

1. People live in which neighborhood likely eat in this kind of restaurant?
2. How many "similar" restaurants are available nearby?
3. Do the "similar" restaurants cost more? If so, what specialty do that have?

Target Audience:

Target audiences for this project does not limit to a person who keeps travelling but everyone. People could simply decide to look for a similar restaurant all the time because they are addicted to a specific category of food. People who rarely use restaurants would prefer to have the most rated restaurants nearby them and all this could be easily handled by our recommender system. So target for this project is basically everyone who is exploring different places or similar places.

2. Data :

To find a solution to the questions and build a recommender model, we need data and lots of data.

Here are some questions we need to consider but not limited to when gathering data to suggest a location for new restaurant:

1. Geographical coordinates (latitude and longitude) to find out where exactly it is located.
2. Population of the neighborhood where the restaurant is located.
3. Average income of neighborhood to know how much is the restaurant worth.

Data Source:

1. Neighborhood data of Toronto from Wikipedia
2. Python geocoder library to get geographical coordinates of neighborhoods
3. Foursquare API venues explore methods to get venues of given neighborhoods
4. Foursquare API venues methods to get ranks and likes of restaurants by given venue id

Methodology:

1. I will use of demographic data by neighborhoods, which allows me to gauge people's taste in each neighborhood. For example, I assume areas with a majority of Chinese people would be prefer to go to Chinese restaurant.
2. I will use the Foursquare API to explore neighborhoods in Toronto and apply the explore function to get the most common venue categories in each of the neighborhoods.
3. I will then use this feature to group the neighborhoods into clusters and use the k-means clustering algorithm to refine the groupings.

4. Finally, I will use the Folium library to visualize the neighborhoods in Toronto and their emerging clusters as well as push the right neighborhood for setting up the restaurant business.