Using clustering technique to recommend Halal restaurant location for opening up

Peiqing Lian (https://github.com/jasperlpq)

1. Introduction

1.1 Background

The purpose of this capstone project is to help people who is looking for a place to open a Halal restaurant in Toronto area. Although, there might have been quite amount Asian food places opening in Toronto, there may not be enough authentic Halal food around and it would be a great opportunity for those who want to start their own restaurant business. Ideas behind this project are divided into two parts, firstly, Halal food, as one of the Asian food branches, is not widely spread around the community. Secondly, there are not a lot of Halal immigration migrating to Toronto and thus helping Halal people to find a great location to start their business would enrich the diversity of the whole community in terms of food and culture.

1.2 Business problem

The objective of this capstone project is to find a trending but suitable location for people who are going to open a Halal restaurant in Toronto area using data science methods such as clustering approach.

1.3 Target audience

The people who want to open up a Halal restaurant in Toronto area

2. Data sources

- We will scrap Toronto neighborhoods data from Wikipedia and corresponding geo information such as latitude and longitude and corresponding borough.

- We the need data of different venues in different neighborhoods of that specific borough.
- In order to gain that information, we will use "Foursquare" locational information. Foursquare is a location data provider with information about all manner of venues and events within an area of interest. Such information includes venue names, locations, menus and even photos. As such, the foursquare location platform will be used as the sole data source since all the stated required information can be obtained through the API. After finding the list of neighborhoods, we then connect to the Foursquare API to gather information about venues inside each and every neighborhood. For each neighborhood, we have chosen the radius to be 100 meters.

The information obtained per venue as follows:

1. Neighborhood
2. Neighborhood Latitude
3. Neighborhood Longitude
4. Venue
5. Name of the venue e.g. the name of a store or restaurant
6. Venue Latitude
7. Venue Longitude
8. Venue Category