Coursera Project: Explore Chinese Restaurant in London

Author: Tzu-Hsuan Chang

In July 2021

Business Problem

- The objective is to help **every tourists , immigrants and long-living foreigners** from East Asian country like China, Korea, Japan, and Taiwan, who seek for the missing flavour or to open up a new business.

- However, which neighbourhoods should be the best decision of opening an Asian restaurant in **London, United Kingdom**?

- How to choose the most suitable places?

- And for tourists, where would be the best neighbourhoods to choose to have cuisine?

DataSets and sources

- **Main Datasets**: Wikipedia link of List of Areas of London

   Including: 1. Borough 2. Locations 3. Latitude 4. Longitude

   BUT coordinates for latitude and longitude would be added on after adjustments with geocoder and create two columns on the dataframe with pandas.

- **Venues Data**: From Foursquare Api, we could acquire the datas from London, United Kingdom for several information of restaurants like categories, different types and styles of cuisines.

Methodology

- Clean the datasets and create a new dataframe with coordinates

- Geocoder mapping to find the locations which would be easier to operate after clustering

- Foursquare API to locate and discover venues of different cuisines

- In each neighbourhoods, using getNearbyVenues() code to acquire the combination data of neighbourhoods, locations, venue, coordinates, and other venue-related data.

Results

Cluster	Most Common Restaurant	2nd Common Restaurant	Least Common Restaurant
1	Italian	South European	Gluten-free
2	Fast Food	Chinese	Halal
3	Depends	Depends	Depends
4	Fast Food	Yoshoku	Halal
5	India	Yoshoku	Greek

- After define the Kmeans and find the clustering model, we define to acquire five Clusters levels

- In the cluster model, we find out that in each neighbourhoods, there are several common restaurants according to the restaurants with Italian, Greek, Chinese, and also gluten-free.

- Sorted by the neighbourhoods and merge the original datasets from Wikipedia and Foursquare. We could find the frequency of restaurants in the new dataframe.

- The result is shown like the form above, with different clusters have each special common restaurants.

Discussion

- Understanding from the output, we could conclude that:

- 1. For owner who would like to open up a new business, suggestion would be made for opening in least common restaurant since there would not be too much competitive.

- 2. For tourists, since we have certain flavour-like restaurants or cuisine, the most common restaurant in each cluster is the main factor to choose as the distinct thought.

Conclusion

In conclusion, this project should have better results since we don not have enough information and data collections to decide several factors like price, races to decide if the populations of each ethnicity are having a decision. But for a general understanding, I think this would be enough, for example, if I would like to open up an Asian-flavoured restaurant, Cluster1 is definitely the best to choose, with the most common restaurants are European restaurants, which means there would not be too competitive.

Thanks for your notice!!!