IBM APPLIED DATA SCIENCE CAPSTONE

Opening a New Chinese Restaurants in Los Angeles, California.

Stephen Kofi Acheampong

June 12, 2020

Outline

Business Problem

- Location of the chinese restaurant is one of the most important decisions that will determine whether the mall will be a success or a failure.
- Objective: To analyze and select the best locations in the city of Los Angeles, California to open a new chinese restaurant.
- This project is timely as a lot of people in the city are currently craving for chinese food.
- **Business question**: In the city of Los Angeles, California, if a chinese restaurant owner or franchisee is looking to open a new chinese restaurant, where would you recommend that they open it?

Data

Data required:

- List of neighbourhoods in the city of Los Angeles, California.
- Latitude and longitude coordinates of the neighbourhoods.
- Venue data, particularly data related to chinese restaurants.

Sources of data

- Wikipedia page for neighbourhoods (https://en.wikipedia.org/wiki/List_of_cities_in_Los_Angeles_County,_California)
- Geocoder package for latitude and longitude coordinates.
- Foursquare API for venue data.

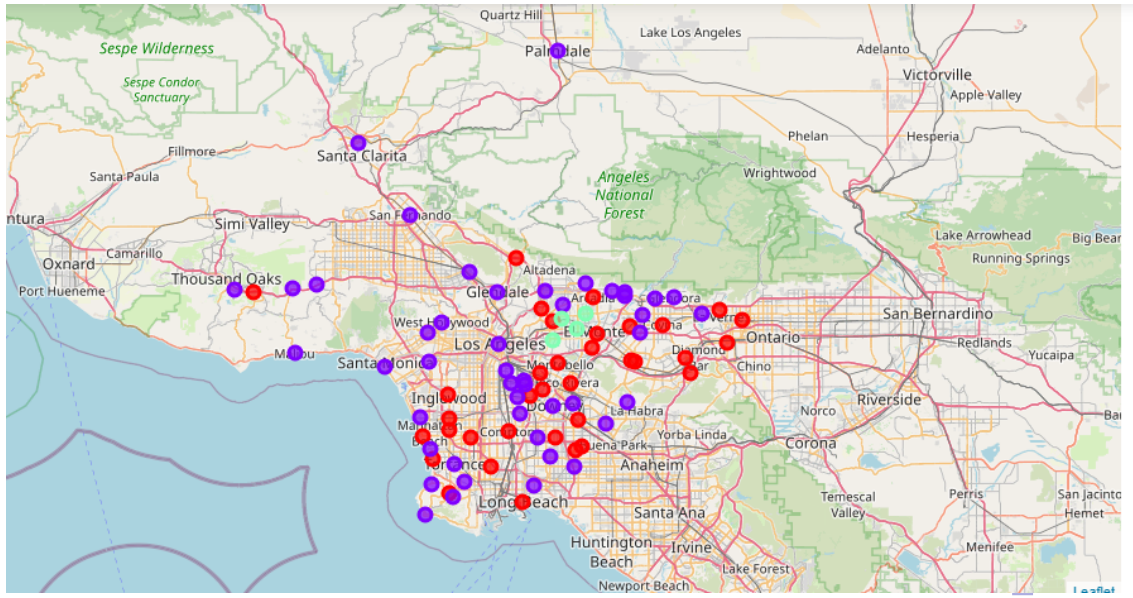
Methodology

- Web scraping Wikipedia page for neighbourhoods list.
- Get latitude and longitude coordinates using Geocoder.
- Use Foursquare API to get venue data.
- Group data by neighbourhood and taking the mean of the frequency of occurrence of each venue category.
- Filter venue category by chinese restaurant
- Perform clustering on the data by using k means clustering
- Visualize the clusters in a map using Folium

Results

Neighbourhoods were clustered into 3 clusters.

- Cluster 0: Neighbourhoods with moderate number of chinese restaurants.
- Cluster 1: Neighbourhoods with the highest number existence of chinese restaurants.
- Cluster 2: Neighbourhoods with the least number existence of chinese restaurants.



Discussion

- Most of the chinese restaurants are concentrated in the central area of the city.
- Highest number in cluster 1 and moderate number in cluster 0.
- Cluster 2 has very low number to chinese restaurants in the neighbourhoods.
- Oversupply of chinese restaurants mostly happened in the central area of the city, with the suburb area still have very few chinese restaurants.
- A differrent number of clusters used instead of 3 in this case may give a different results.

Recommendations

- Open new chinese restaurants in neighbourhoods in cluster 2 with little competition.
- One may also open in neighbourhoods in cluster 0 with moderate competition if have unique selling propositions to stand out from the competition.
- Cluster 2 has very low number to chinese restaurants in the neighbourhoods.
- Avoid neighbourhoods in cluster 1, already high concentration of chinese restaurants and intense competition.

Conclusion

- Answer to business question: The neighbourhoods in cluster 2 are the most preferred locations to open a new chinese restaurant.
- Findings of this project will help the relevant stakeholders to capitalize on the opportunities on high potential locations while avoiding overcrowded areas in their decisions to open a new chinese restaurant.

THANK YOU