COURSERA CAPSTONE

IBM APPLIED DATA SCIENCE CAPSTONE

OPENING A NEW CHINESE RESTAURANT IN SAN DIEGO, CALIFORNIA

BY: SHAUN CLARKE

NOVEMBER 2019

Business Problem

- Location of the Chinese Restaurant is one of the most important decisions that will determine whether the restaurant will be a success or a failure

- Objective: To analyze and select the best locations in San Diego, CA to open a new Chinese Restaurant

- This project is timely as the city is currently suffering from oversupply of Chinese Restaurants

- Business question
  - In San Diego, if a property developer is looking to open a new Chinese Restaurant, where would you recommend that they open it?
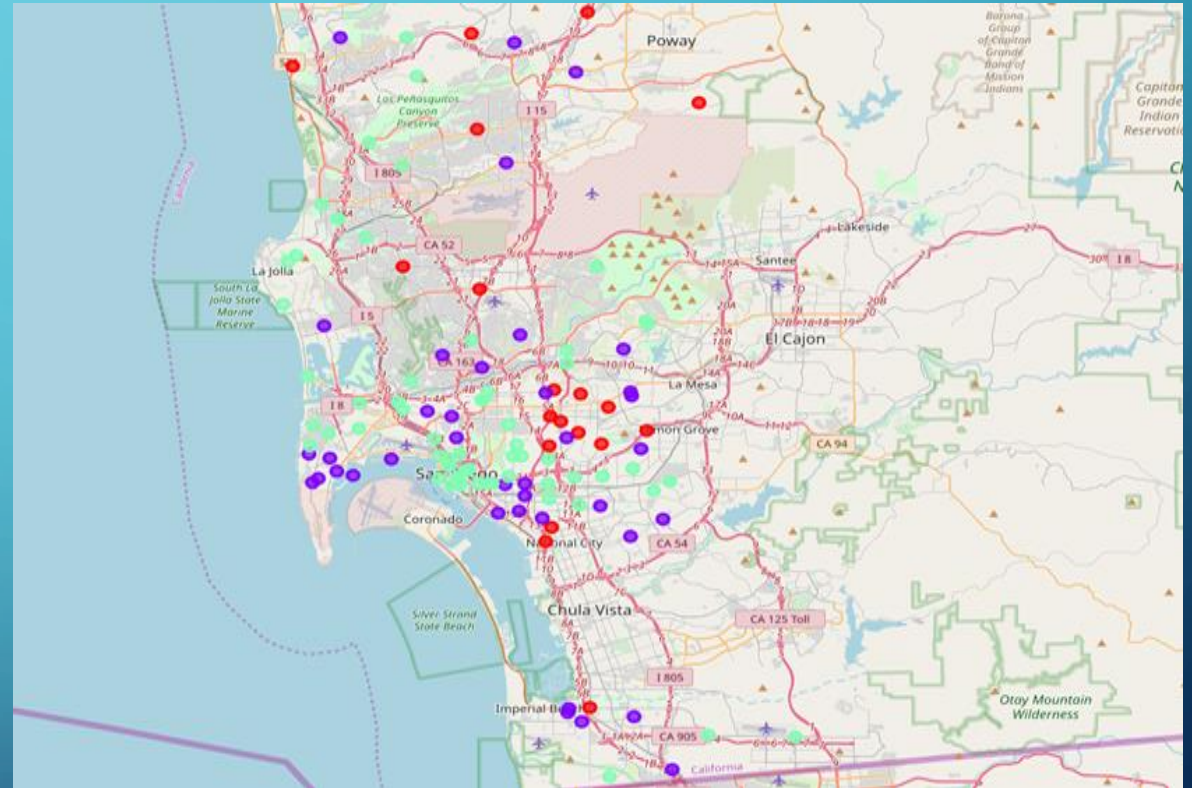
Data

- Data required
  - List of neighbourhoods in San Diego
  - 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/Category:Neighborhoods_in_San_Diego)
  - Geocoder package for latitude and longitude coordinates
  - Foursquare API for venue data

Methodology

- Web scraping Wikipedia page for neighborhoods list

- Get latitude and longitude coordinates using Geocoder

- Use Foursquare API to get venue data

- Group data by neighborhood and taking the mean of the frequency of occurrence of each venue category

- Filter venue category by Chinese restaurants

- Perform clustering on the data by using k-means clustering

- Visualize the clusters in a map using Folium

Results

- Categorized the neighborhoods into 3 clusters:
  - ❑ Cluster 0: Neighborhoods with low number to no existence of Chinese Restaurants
  - ❑ Cluster 1: Neighborhoods with moderate number of Chinese Restaurants
  - ❑ Cluster 2: Neighborhoods with high concentration of Chinese Restaurants

Discussion

- Most of the Chinese restaurants are concentrated in the central area of San Diego

- Highest number is in cluster 2 and moderate number in cluster 1

- Cluster 0 has very low number to no Chinese Restaurants in the neighborhoods

- Over saturation of Chinese restaurants are mostly visible in the central area of the city, with the suburbs having only a few

Recommendations

- Open new Chinese Restaurants in neighborhoods in cluster 0 with little to no competition

- We can also open in neighborhoods in cluster 1 with moderate competition along with a unique selling point to stand out from the competition

- Avoid neighborhoods in cluster 2, they already have a high concentration of Chinese Restaurants and intense competition

Conclusion

- Answer to business question: The neighborhoods in cluster 0 are the most preferred locations to open a new Chinese Restaurant

- The findings of this project will help the relevant stakeholders to capitalize on locations with a high potential for success. while avoiding overcrowded areas in their decisions to open a new Chinese Restaurant