Capstone Project
The Battle of Neighbourhoods

By Parthasarathi Das (parthasdas@hotmail.com)
Date : 08-Apr-2020

Contents

1. Introduction

1.1   Background

In order to open a business venture it is important to carry out some feasibility analysis regarding the potential. Especially in a business that has low entry barrier (e.g Restaurant), it is better to carry out some ground analysis that covers Competition among other factors.

1.2   Problem

A friend is looking to open a restaurant. He wants to avoid big cities like New York, San Francisco.

1.3   Interest

He is looking for some relatively small cities in USA and wants a analysis report on the feasibility.

2. Data Acquisition and Data Preparation

Data had to be collected from multiple sources and stitched together to do the analysis and recommendation.

First, I searched the Internet to get a listing of the cities and their population. This was available in Wikipedia link mentioned below.

https://en.wikipedia.org/wiki/List_of_United_States_cities_by_population

The census data was from 2018 but that was a close match.

Foursquare API was used to gather the information about venues in the two cities.

3. Methodology

The Wikipedia site lists the 314 incorporated places in the United States with a population of at least 100,000 as on July 1, 2018, as estimated by the **United States Census Bureau**. Five states—Delaware, Maine, Vermont, West Virginia and Wyoming—have no cities with populations of 100,000 or more.

The technique of screen-scrapping was used read the city information into a panda DataFrame. It was found that two cities namely Chicago and Houston had population in the range of 2.5million. These two cities were picked up as possible choices for exploring the potential of opening a new restaurant. As of 2018 estimate, Chicago had a population of 2.7 Million while Houston had a population of 2.3 Million.

Population constitute the most important factor for a Restaurant. There are other factors such as City culture, Per-capital income, ethnicity and even climate. But the remaining are secondary.

The second data source was the Four-square API. Using the Explore option, the venues were read into panda DataFrames. While invoking the 'Explore' API for both the cities, a radius of 750 meters were taken and the category of "food" was taken. This filtered out other venues like Museums, theatres etc. So now we had the list of food joints of the two cities along with the categories. We got a total list of 588 food venues for the two cities

Next, from the list of the food venues, the categories of Bistro, cafes, joints, food-truck were filtered out. So now we had a clean list of pure-play restaurants in the two cities.

The next step was to generate a pivot table to give a summary of the restaurants type for the two cities in a format for easy comparison.

	Venue	
Neighborhood	Chicago	Houston
Venue Category		
American Restaurant	2	0
Cantonese Restaurant	1	0
Chinese Restaurant	2	0
Dumpling Restaurant	0	1
Empanada Restaurant	0	1
Fast Food Restaurant	1	0
Greek Restaurant	0	1
Italian Restaurant	1	5
Japanese Restaurant	0	1
Latin American Restaurant	1	0
Mexican Restaurant	6	4
New American Restaurant	0	1
Seafood Restaurant	0	1
Southern / Soul Food Restaurant	0	2
Steakhouse	0	2
Sushi Restaurant	1	1
Szechuan Restaurant	1	0
Taiwanese Restaurant	1	0
Thai Restaurant	0	1
Vietnamese Restaurant	2	0

4. Results

From the above tabulation, the observations are as below.

a. It appears that Mexican Restaurants are quite numerous in both the locations (6 and 5). Opening a Mexican Restaurant in either city is going to face competition
b. There are 5 Italian restaurants in Houston but only 1 in Chicago. There is a potential business opportunity
c. Chicago does not seem to have a steakhouse at all. That should be a business opportunity here.

¶

5. Observations & Recommendations

Overall, recommendation should be to go for a Italian Restaurant or a steakhouse in Chicago.

Now, as far as the location is concerned, a map-plotting using Folium shows that the restaurants are concentrated on two particular stretches of the city. In other words, they

for two clusters. It is advisable that the new restaurant be opened in one of the clusters. It might **seemingly** be counter-productive but there is a rationale to this. Citizens of Chicago would have a perception or mind-share about the two clusters even if they are not formally identified/indicated. On a typical holiday or weekends, people wanting to dine-out would head for these clusters keeping the option open for a type of cuisine/restaurant. Having the new outlet there would have a potential footfall from the crowd thronging to those areas.
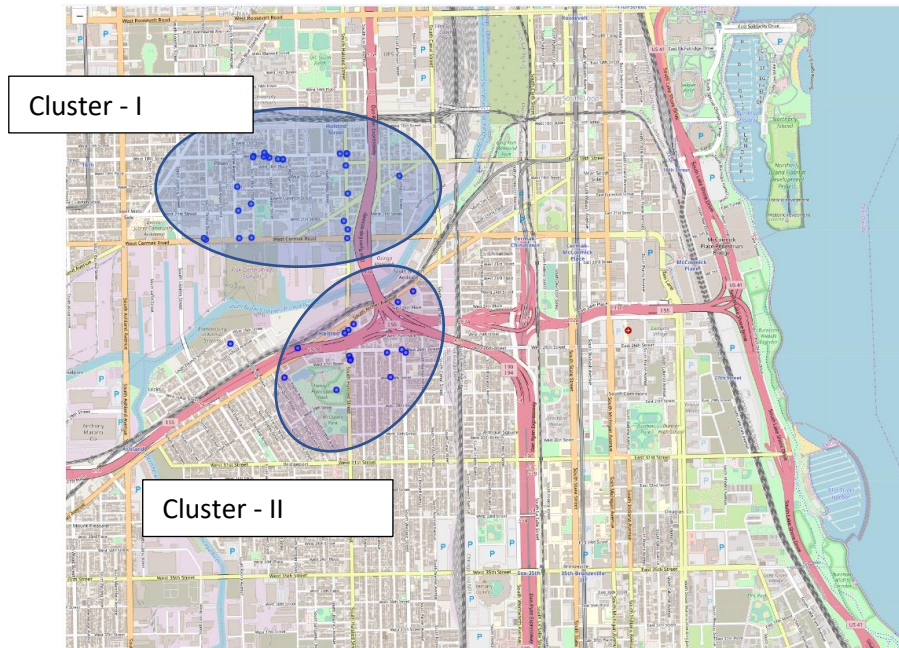
6. Conclusion

It needs to be mentioned however, that there are two factors which should have been included in the feasibility analysis. These are i) Per-capita income and ii) Ethnicity of the population(% distribution and not just the names). These data points were not readily available and hence these factors could not be included in the analysis.

Needless to say, higher the per-capita income , more the money at disposal and better the chance of a restaurant to attract patrons.

Ethnicity also drives the footfall. While nothing as such prevent a person of Greek origin to visit a Russian restaurant, it is generally observed that a person living away from his /her origin will have a sentimental attachment to culture and cuisine that he left behind. The sentimental attachment will drive him to the restaurant tied with his origin.

Hopefully, future research in this aspect will include the two factors mentioned.