OPENING A CHINESE RESTAURANT IN NEW YORK

I. Introduction

1.1.Background

New York City has long been one of the most sought-after places to open a restaurant. It's also the most populous city in the US with millions people and counting; that's a lot of hungry mouths to feed, so it's no wonder more and more restaurant owners are setting up shop. Data from the National Restaurant Association shows that there are almost 45,000 food establishments in New York City, so restaurant entrepreneurs can expect stiff competition. But there's still room for business growth.

Each borough of New York City has a completely different vibe, which means your concept might perform great in one of them, but terrible in one or more of the others:

- Manhattan is the borough that comes to most people's mind when they think of New York City—Times Square, Central Park, the Empire State building. While you can find almost any type of restaurant that you would want in Manhattan, this is the most expensive borough to both live and do business in. For a new restaurant owner with limited capital, the high cost of real estate and intense competition are both reasons to consider an area outside of Manhattan;
- The Bronx offers a cheaper alternative to Manhattan and still features tourist attractions like The Bronx Zoo, Yankee Stadium, and Little Italy. Plus, schools like Fordham University and Manhattan College means there are plenty of hungry students looking for a place to eat. If you're just starting out in New York City, The Bronx could offer an affordable place to start.

- Brooklyn is quickly becoming the “it” spot for hip, young millennial. If this is the target audience you’re trying to attract, then Brooklyn is the place to be. But keep in mind, while this area is becoming increasingly popular, you can expect your operating costs to increase as well.
- Queens is home to New York City’s largest borough and houses nearly two million people. Astoria is one of the most well-known neighborhoods in Queens and boasts beautiful parks, plenty of restaurants, and nightlife all while being a quick subway ride away from Manhattan. This could be an ideal area for restaurants looking to cater to young families and professionals.
- Staten Island is often New York City’s overlooked borough, but is home to several historical sites, museums, and botanical gardens and offers another less expensive alternative to Manhattan and Brooklyn.

1.2. Business problem

In this project, I will try to find an optimal location for a restaurant. The project aims to **answer the question**: if someone is looking for a place to open a new Chinese restaurant in New York, where would he/she take as a recommendation?

1.3. Target audience

Specifically, this report will be targeted to stakeholders interested in opening a Chinese restaurant in New York, US.

II. Data acquisition and cleaning

2.1. Required factors

Based on definition of our problem, factors that will influence the decision are:

1. Number of existing restaurants in the neighborhood (any type of restaurant);

2. Number of and distance to Chinese restaurants in the neighborhood, if any distance of neighborhood from city center.

(Use regularly spaced grid of locations, centered on city center, to define our neighborhoods.)

2.2. Data sources

- Map of New York and its boroughs (got json file from NYC Open Data https://opendata.cityofnewyork.us/)
- Coordinates of NY center (I chose Queens Center since it's in the middle of the map), coordinates of each borough (got data from LocationIQ API)
- Venue data: restaurants in general and Chinese ones in New York (got data from Foursquare API)

2.3. Methods to generate data

Data sources will be generated the required information:

- Centers of candidate areas will be generated algorithmically and approximate addresses of centers of those areas will be obtained using LocationIQ API reverse geocoding;
- Number of restaurants and their type and location in every neighborhood will be obtained using Foursquare API;
- Coordinate of The Bronx center will be obtained using LocationIQ API geocoding of a well-known location (Queens Center)

III. Methodology

In this project, we will try to detect areas of New York that have low restaurant density, particularly those with low number of Chinese restaurants by limiting our analysis to area 25km around city center.

Firstly, we need the location and type (category) of every restaurant within 25km from NY center (Queens Center). Also, we have to identify the Chinese restaurants according to Foursquare categorization.

Secondly, after using the Foursquare API to make the call, we will have the number of restaurants in NY and number of Chinese restaurants there. We will also have each venue's name, latitude and longitude. With the data, we will calculate and explore the 'restaurant density' across different areas of New York by using heat maps, and then identify a few promising areas close to center with low number of restaurants in general and focus our attention on those areas.

Finally, we will focus on most promising areas, and within those create clusters of locations that meet some basic requirements established in discussion with stakeholders: we will take into consideration locations with no more than two restaurants in radius of 1000 meters, and we want locations without Chinese restaurants in radius of 1500 meters. We will present map of all such locations but also create clusters **using k-means clustering** of those locations to identify general zones/neighborhoods/addresses which should be a starting point for final 'street level' exploration and search for optimal venue location by stakeholders.

IV. Results and discussion

The analysis shows that although there are a great number of restaurants in NY (2124 in our initial area of interest); there are pockets of low restaurant density fairly close to city center. Highest concentration of restaurants was detected north and west from Queens Center, so we focused focus our analysis on **north-east, south and south-west** from New York center, corresponding to boroughs Brooklyn, Queens and north-east corner of Bronx.

Although another borough was identified as potentially interesting (Staten Island), our attention was focused on Brooklyn and Queens, which offer a combination of popularity among tourists, closeness to city center, strong socio-economic dynamics and a number of pockets of low restaurant density.

After directing our attention to this more narrow area of interest, we first created a dense grid of location candidates. Those location candidates were then clustered to create zones of interest which contain greatest number of location candidates. Addresses of centers of those zones were also generated using reverse geocoding to be used as markers/starting points for more detailed local analysis based on other factors.

Result of all this is 15 zones containing largest number of potential new restaurant locations based on number of and distance to existing venues - both restaurants in general and Chinese restaurants particularly. This does not imply that those zones are actually optimal locations for a new restaurant; the purpose of this analysis was to only provide info on areas close to NY center but not crowded with existing restaurants (particularly Chinese). It is entirely possible that there is a very good reason for small number of restaurants in any of those areas, reasons which would make them unsuitable for a new restaurant regardless of lack of competition in the area. Recommended zones should therefore be considered only as a starting point for more detailed analysis which

could eventually result in location which has not only no nearby competition but also other factors taken into account and all other relevant conditions met.

V. Conclusion

Purpose of this project was to identify NY areas close to center with low number of restaurants (particularly Chinese restaurants) in order to aid stakeholders in narrowing down the search for optimal location for a new Chinese restaurant. By calculating restaurant density distribution from Foursquare data, we have first identified general boroughs that justify further analysis (Brooklyn and Queens), and then generated extensive collection of locations which satisfy some basic requirements regarding existing nearby restaurants. Clustering of those locations was then performed in order to create major zones of interest (containing greatest number of potential locations) and addresses of those zone centers were created to be used as starting points for final exploration by stakeholders.

Final decision on optimal restaurant location will be made by stakeholders based on specific characteristics of neighborhoods and locations in every recommended zone, taking into consideration additional factors like attractiveness of each location (proximity to park or water), levels of noise / proximity to major roads, real estate availability, prices, social and economic dynamics of every neighborhood etc.