**Launch a Chinese Restaurant in LA County**

Gutuan Zheng

**Introduction**

To open a restaurant, location is the No.1 question to ask. It is even more important to open a Chinese restaurant in the Los Andreas area, where the ethical demography is broadly spreading-out in the western suburban.

Location is important for a simple reason, reducing the fatality risk. The restaurant business has a high failure rate. Around 60 percent of new restaurants failed within the first year. And nearly 80 percent closed door before their fifth anniversary.

We need to identify the best location where optimized conditions exist to sustain a healthy business. There are many factors to consider, such as sufficient parking space, accessibility, visibility, etc…In this study, to gain perspective into the location choice, we will focus on only two metrics we think effect most:

* Ethical Population: There shall be a plural Chines-Americans living in the nearby communities to maintain the foot traffics.
* Competition: There shall not be too many niche restaurants operating in the same area, so enough revenue can be expected to offset fixed cost and realize sizable profit.

1. **Business Problem**

We are commissioned to identify the best location to open a Chinese restaurant in the Los Angeles area. The community shall have a high Chinese American population, and the Chinese cuisine establishment competition is modulated. Let us look into those two issues in more details.

**1.1 Chinese-American populations in Los Angeles Area**

Los Angeles is one of the three [metropolitan areas](https://en.wikipedia.org/wiki/Metropolitan_Statistical_Area) with the largest Chinese-American populations, according to the recent Census estimates.

The [Greater New York](https://en.wikipedia.org/wiki/New_York_metropolitan_area) [Combined Statistical Area](https://en.wikipedia.org/wiki/Combined_Statistical_Area) has 735,019 Chinese-American people, the [San Jose-San Francisco-Oakland](https://en.wikipedia.org/wiki/San_Francisco_Bay_Area) Combined Statistical Area has 629,243 Chinese-American people, and the [Greater Los Angeles](https://en.wikipedia.org/wiki/Greater_Los_Angeles_area) Combined Statistical Area has about 566,968 Chinese-American people. In the post-1965 era, first- and second-generation immigrants include those from [Mainland China](https://en.wikipedia.org/wiki/Mainland_China), [Hong Kong](https://en.wikipedia.org/wiki/Hong_Kong), [Macau](https://en.wikipedia.org/wiki/Macau), and [Taiwan](https://en.wikipedia.org/wiki/Taiwan). Also included in the Chinese-American population enumeration are [ethnic Chinese from Malaysia](https://en.wikipedia.org/wiki/Malaysian_Chinese) and [Vietnam](https://en.wikipedia.org/wiki/Hoa_people) who might identify themselves as Chinese.

The Los Angeles County city of [Acadia](https://en.wikipedia.org/wiki/Monterey_Park,_California), with 26,229 Chinese American, has the highest percentage of Chinese-Americans of any municipality, at 45.6% of its total population. The [San Gabriel Valley](https://en.wikipedia.org/wiki/San_Gabriel_Valley) region of [Los Angeles County](https://en.wikipedia.org/wiki/Los_Angeles_County) is the single largest concentration of combined Chinese and Taiwanese Americans in the country, having a collections of U.S. suburbs with large foreign-born Chinese-speaking populations, ranging from working-class individuals residing in [Rosemead](https://en.wikipedia.org/wiki/Rosemead,_CA) and [El Monte](https://en.wikipedia.org/wiki/El_Monte,_CA) to wealthier immigrants living in [Arcadia](https://en.wikipedia.org/wiki/Arcadia,_CA), [San Marino](https://en.wikipedia.org/wiki/San_Marino,_CA), and [Diamond Bar](https://en.wikipedia.org/wiki/Diamond_Bar,_CA), reflecting their general affluence and propensity for [professional](https://en.wikipedia.org/wiki/Professional) occupations.

The majority of the Chinese-American population of more than 400,000 in [Los Angeles County](https://en.wikipedia.org/wiki/Los_Angeles_County) lives within the [San Gabriel Valley](https://en.wikipedia.org/wiki/San_Gabriel_Valley), which is particularly pronounced for cities that have a large Chinese-American ethnic plurality. The following cities have the highest percentage of Chinese-Americans in [Greater Los Angeles](https://en.wikipedia.org/wiki/Greater_Los_Angeles_Area). Those cities are the potential candidates to launch a Chinese restaurant business.

*Table 1: Chinese-American Population by City in LA (Source: Wikimedia)*



**1.2 Chinese Restaurant Business in Los Angeles Area**

The Chinese Restaurant Business in Los Angeles Area can be characterized in the following two ways: (1). More variety of Chinese dishes are served among fierce competition; (2). Upscale new restaurants established recently to attract the new affluent generation of Chinese Americans as well as non-Chinese ethical clients.

**Variety:** Los Angeles, with one of the largest ethical Chinese populations outside Asia, is widely regarded as an epicenter for Chinese cuisine. Although the best Chinese restaurants are mostly concentrated in the ethnic enclaves of the San Gabriel Valley, there are increasingly more in neighborhoods throughout L.A., from historic Chinatown to the Westside.

Los Angeles has the most diverse population of Chinese, which makes all the difference in the dining scene. With the one-menu-feed-all style “American Chinese” Chinese restaurants still going strong, newly established Chinese restaurants are more focusing on the differentiations. The Los Angeles Chinese food scene is incredibly diverse. Of the 34 official provincial-level administrative units in China, 21 of them are represented in restaurants in the greater Los Angeles area. Some offers the foods mainly found in the north of China with tons of wheat, while others serve Southerners mainly with rice. One can find not only the spicy food in Chinese southwest oriented restaurant, but also the moderated tastes only slightly using chili pepper.

**Upscale**: For a very long time the Chinese food in America is the synonymous of cheap food with fried noodles and sticky-sweet meat nuggets. People thought the Chinese food is quick and greasy, and shouldn’t cost very much. Not any more. The new generation of Asian American chefs in California today are challenging the long-standing stereotype, and coming up with new looks and new tastes, as a result, new value. Other than the famous P F Chang, many upscale and fancy Chinese restaurants are established. Actually, the No.1 ranked Chinese restaurant in the US is Chengdu Taste, as expected, located in LA. The evolution is not only limited to the food itself, but also other elements typically absent from traditional Chinese restaurants — tasting menus, service emphasis, luxe decor, innovative cocktail menus and wine pairings — all are showing up in the upscale Chinese restaurants. This new shift can be found in the LA downtown as well as the San Gabriel Valley westward. The change is not a simply Western focus, but more about the future of the Chinese cuisine. This young generation of chefs is combining childhood nostalgia with professional culinary chops, local ingredients and fine dining mores. Those new upscale Chinese restaurants propel the perception and acceptance of Chinese cooking into new directions and into new neighborhoods, generating new markets.

As we can tell from above discussion, Chinese Restaurants are facing fierce competitions. To gain an edge, one have to differentiate. The business owners are innovating to break the mold to survive and make a profit. Location selection becoming more important to alleviate the risks.

1. **Data Collection**

To search for the best location to launch a new Chinese restaurant in the Los Angeles area, we need to collect two sets of data: The Chinese American distribution by the City, and the Chinese and Asian restaurant spreads by the City.

Data quality and reliability are top criteria in our information search. In case the data come from the websites, we make sure they also sourced from authorities to ensure accuracy of the data.

**2.1 Top Cities with Chinese-American in Los Angeles Area**

Ethical population data can be found in the national census web site as well as the institutions studying the ethical developments. Wikimedia is another source to summarize the population migration and growth. We have examined all above sources.

The national census web site (https://www.census.gov/) contains most reliable and detail population data for in-depth study. One can search for population distribution by age, ethics, region, immigration origin, etc. The tools on that website is very useful, especially the “Data via FTP” tool, making it convenient to retrieve data. Even though, depending the question, one may have to check into many data tables to put together a full set of information.

Immigration focused agencies such as MPI (Migration Policy Institute) can provide many meaningful insights into the ethical issues. They also collect data in their study areas, and summarize the information to provide guidance to policy makers. With MPI as an example, they study the i[mmigration over time](https://www.migrationpolicy.org/programs/data-hub/us-immigration-trends#history), [immigrants' countries and regions of birth](https://www.migrationpolicy.org/programs/data-hub/us-immigration-trends#source), [educational attainment](https://www.migrationpolicy.org/programs/data-hub/us-immigration-trends#education), etc….Had one found relevant study articles, one can quickly gain deep insights into the problems and seek for possible solutions.

Unfortunately, for our case, where we need to derive the data by city for a specific ethics, above sites have not delivered handy details to advance our study. We further searched the Wikimedia, and fortunately found out the exact data with Chinese American population by City in the Los Angeles area, as shown in the table 1 in the 2.2 section of this article. Not only the population is presented, but also the penetration rate is listed for added insights.

It shows that the majority of the Chinese-American population of more than 400,000 in [Los Angeles County](https://en.wikipedia.org/wiki/Los_Angeles_County) lives within the [San Gabriel Valley](https://en.wikipedia.org/wiki/San_Gabriel_Valley), which is particularly noted for cities that have a large Chinese-American ethnic plurality. There are 5 cities having Chinese American population exceeds 40% of total population, including Arcadia (46%), Monterey Park (45%), Temple City (44%), San Marino (43%), and San Gabriel (42%). Alhambra has most population at 30,683 Chinese American. All those cities are potential locations for our new business.

**2.2 Chinese and Asian Restaurant Business Distribution**

We leverage the Foursquare to acquire the Chinese and Asian restaurant information. Foursquare is a location technology platform dedicated to provide information for people moving through the real world. It is the location engine supporting many mobile apps to allow people to conduct data-driven media planning. The venue based search can help people quickly drill into the targeting dataset.

We will need the location information, such as the latitude and longitude, to execute the Foursquare searches. Geopy library will be used to retrieve the data. More specifically, we use the geolocator to derive data. Just we have to be a bit patient in case it “times out”.

With the pre-acquired Client ID and Client Secret codes from Foursquare, I can leverage the geolocator-derived location data to retrieve the top venue info through Foursquare for each of the top cities in Los Angeles area, where high concentration of Chinese Americans present.

Data processing will be needed to cleanse data. Out of 185 venue obtained, we first group the top venue categories for the selected cities. Attention is especially paid to the Chinese restaurant and other Asian restaurants, for the latter is also a key factor affecting our location decision. The final feature list include in the study are: Chinese restaurant, Korean restaurant, Sushi restaurant, Vietnam restaurant, and Indian restaurant. Based on those Asian restaurant spreads, one can use the cluster technique to group the Cities into categories to narrow down our location selection.

We use the Geo visualization to help further fine tune the location decision. We will avoid the areas that having multiple neighboring cities with high density of Asian restaurant, and look into other factors, such as overall population, etc… to identify the optimized city.

By combining the ethical population spread and competitive business distribution, we will derive a shortlist of cities in the Los Angeles area to launch a new Chinese restaurant business.

**3. Methodology**

We examine the factors that will affect the decision making for the preferred location to launch a Chinese restaurant business, with the targets of maximizing the client foot traffics and minimizing the competition.

To secure a sufficient cliental we primarily looked into the cities in the Los Angela’s area that have a high level of ethical population. Additional considerations such as accessibility, parking space, …will be included only after we narrowed our selection to a short list of cities. Ethical population information is sourced from Wikimedia, augmented with studies from influential think tanks, as well as the additional data from the national census bureau.

The reading of the competing shop distributions in this area is retrieved by Foursquare, from which we can scrape the data to investigate the density of Asian restaurants by the city.

The analysis is conducted after marring the two sets of data. Specifically, we examine the data in the following ways:

* Identify the cluster of cities that do not have any Asian restaurants, or have only a few ones. By doing so, we exclude the area where the business had been presented, or even saturated.
* Visualize the cluster of cities per the Asian restaurant features, and further exclude those cities that does not have Asian restaurants, but in the close proximity to the high density of competing business.

We employed several data science techniques to carry out the study, including the Geolocator to derive the location information, K-Cluster to group the cities based on relevant features, and Folium to plot the Los Anlage’s area analytical information. Foursquare platform is employed to acquire the aggregated city venue data.

**3.1. LA Community Chinese Population Data**

We scraped the Chinese American population data by city from a Wikipedia webpage. Because that webpage contains multiple tables in various format, we utilized the Pandas and lxml to pin point the table (7th) we need to retrieve, and pulled the data into a dataframe for our analysis. That data set contains the top 42 cities with highest penetration of Chinese American, and also include all top cities with high volume of ethical populations.

**3.2. Retrieve Geo Location Data**

Geolocator is used to calculate the latitude and longitude of all cities in our study. The challenge had been the timing out after too many runs. We adopted two means to make sure we will have the full set of data: (1).Added a timer to pause the program by 1 second after each city run, so that the server burden to be alleviated; (2). Save data into an external csv document once retrieved, so that we retrieve data only once during all iterations of data testing. Be cautioned that the data precision accuracy will be reduced if we just do a simple to\_csv. It is suggested to add a float data format when saving the data to preserve the decimals.

**3.3. Retrieve Geo Venue Data via Foursquare**

The following set of information are required to retrieve data from Foursquare

* Credentials (Client ID and Client Secret Code)
* Version of Foursquare data (“20180605” used)
* Geo coordinates to pull data
* Radius specifying the area size data to be pulled from (2000, cover downtown area but not overlap nearby cities)
* Limit, the number of trial to pull data (100)

Venue data are obtained through a two-step approach: (1). Understand the top venues in this area through a typical city of choice, such as Acadia; (2). Parsing through all top cities to acquire all venues for all cities.

**3.4. K-Cluster Analysis**

The venues from the all city runs are grouped and analyzed using the K-Cluster technique. We expect the clustering will help us narrow down the sections. Since the location decision can be affected by many factors, too many number of cluster is not recommended. With a well controlled set of features, we select K=3 for our study.

We noticed that there are totally 185 different venues, such as ATM, Art Gallery, Trail, fast Food, etc….from Foursquare data. For our purpose, we like to select only those that related to the competition to a Chinese restaurant, i.e., the Asian food shops, including Chinese restaurant, Korean restaurant, Sushi restaurant, Vietnam restaurant, and Indian restaurant.

Actually, we first run a K-cluster analysis of all cities based on all the 185 features. The result showing a quite homogenous distribution, without much outstanding characteristics. This is expected, because all of them are Chinese American concentrated cities, with a lot of similarity per their choice.

The second run of K-Cluster is based only on 5 features of our selection, focusing on relevant Asian food shops. The result is now polarized, reveal the cities without few Asian restaurant vs those have multiple.

**3.5. Visualization of Clusters**

More factors need to be evaluated other than the number of competing business in a city. One consideration is the neighboring cities that already occupied by high volume competing business. To better assess the situation, we visualize the clustering in a map, to assist further judgement calls.

**3.6. Plot Comparison per Population and Number of Business**

The analysis can be further verified by plotting the Chinese American population and the number of relevant business in a single plot, making the decision further visualized for final decision. We have been examine the data in various ways to understand the data structures.

1. **Result**

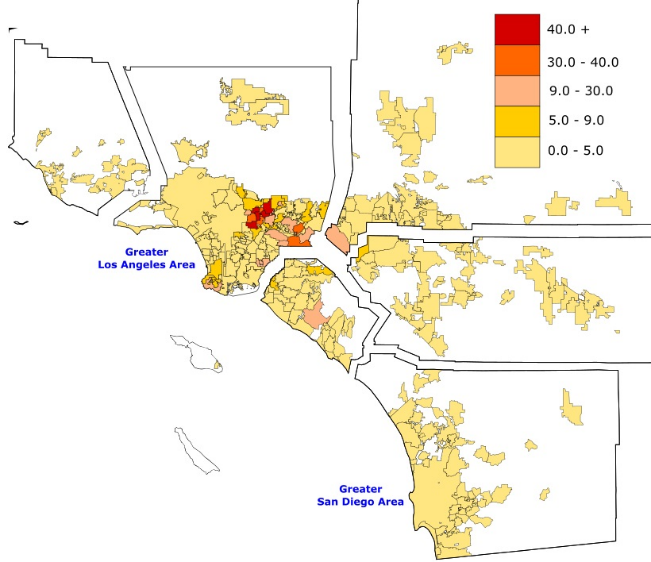
To launch a Chinese restaurant business, we sourced data from Wikimedia and Foursquare to narrow down the list of cities in the Los Angels area. Staring from all cities in the Los Angeles county, we employ ML technique to filter out the cities has less desirable business environment. Step by step we drill into the data to derive our short list of candidates.

**4.1 Top 42 Cities with High Level Population of Ethical Population**

Of all the 88 cities in Los Angeles county, we initially cut the number by about half, through selection of the 42 cities with high population of Chinese Americans. Top candidate for our business include Arcadia (46%), Monterey Park (45%), Temple City (44%), San Marino (43%), and San Gabriel (42%). Alhambra has most population at 30,683 Chinese American. All other cities in the list also demonstrated high level ethical population and high level ethical penetrations, which is essential to support an ethical flavored restaurant.

As shown in the map. Most of those 42 cities are in the eastern suburban of Los Angeles. Majority in San Gabriel Valley in Northwestern, while a few in southeastern are. A couple outliers situated near the seashore, and into the mountain area far east.

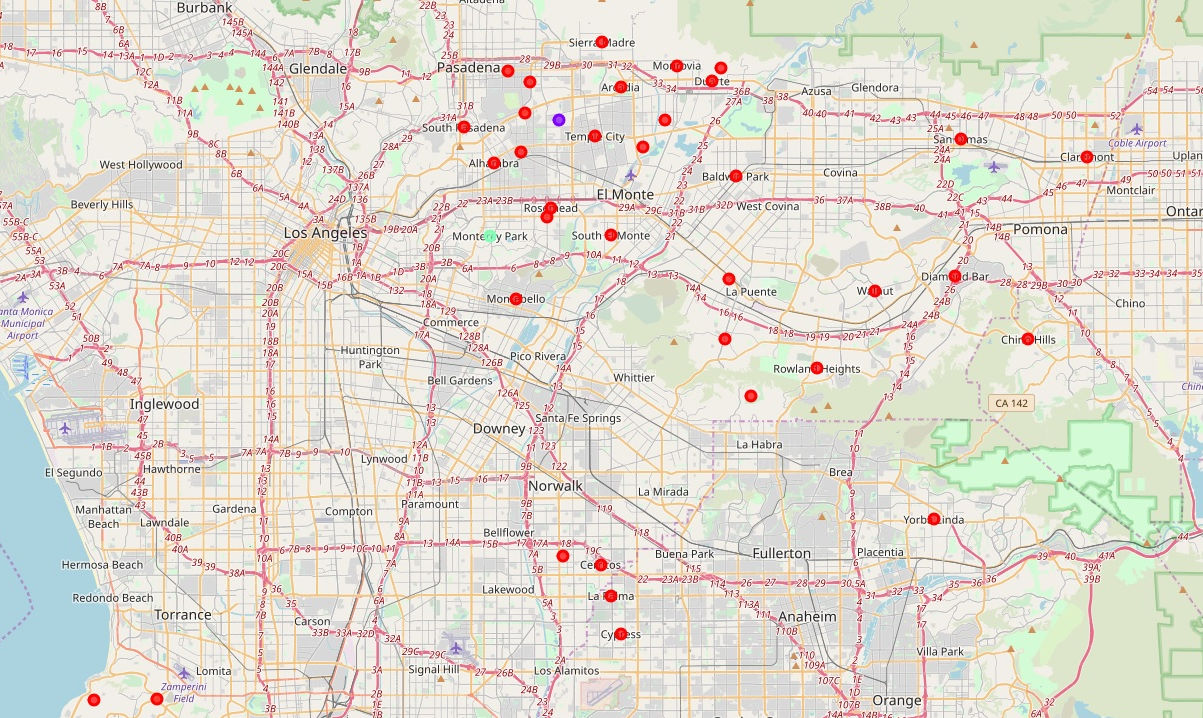
*Figure 1. Los Angeles Area Chinese American Distribution*



**4.2 City Clustering Based on All Venues**

We further examine the business landscape in those cities. First we retrieve the top venues in all the cities, and utilized all the venues found in those cities to cluster them. As seen from the map shown below, the cities are quite homogeneous, as expected, for the similarities the Chinese Americans select to settle.

*Figure 2. City Clusters per All Venues in Los Angeles County*



**4.3 City Clustering Based on Asian Restaurants**

We re-run the clustering analysis, this time, based on only the Asian restaurants, which we think are the key competing parameters that will affect our decisions. Interesting enough, we identified 4 cities that are highly competitive in the Asian restaurant business, that we would like to avoid.

La Palma city singled out as a unique city cluster dominating the southeastern area of Los Angeles.

Another cluster all featured by density operated Asian food shops in the northeastern area of Los Angeles. Those 4 cities are further excluded from our candidate list.

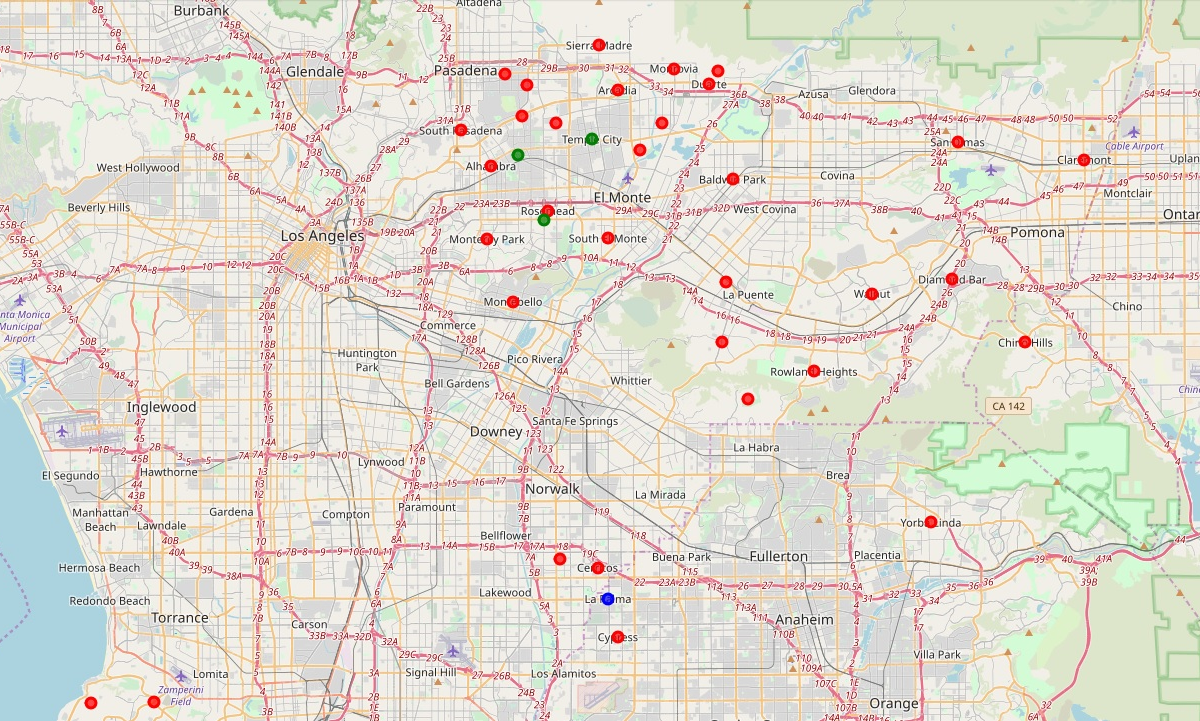
*Table 2: Clusters based on All Venues*

|  |  |  |
| --- | --- | --- |
| Cluster Order | Count of Cities | City List |
| 0 | 37 | Acadia, Monterey Park, San Marino… |
| 1 | 3 | Temple City, San Gabriel, South San Bagriel |
| 2 | 1 | La Palma |

**4.4 Exclude Cities In Proximity of High Competitive Cities**

We plot the clustering information in the Los Angeles area map, and noticed that many cities, though have no or few Asian restaurant operations, but located very close to the cities that do have high number of Asian food business.

*Figure 3. City Clusters per Asian Foods in Los Angeles County*

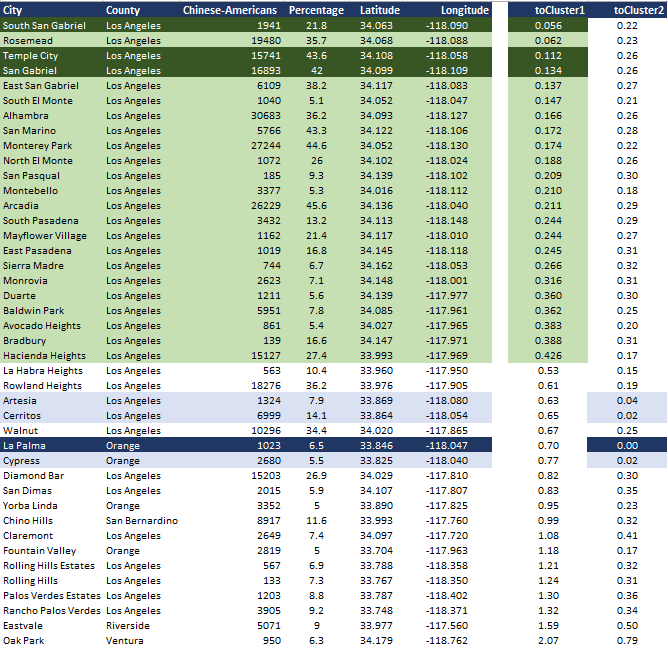


La Palma serves not only itself, but also covered three other cities near it, including Artesia, Cerritos, and Cypress. Given the proximity, we will group those three additional cities into generalized cluster 2, with 1+ 4 = 4 cities.

For cluster 1, we also need to include those cities with a short driving distance from the three cluster 1 cities. To do that, we calculate the distance of other cities to these three cities, then add them together. After sorting all the cities, we identified additional 19 cities that in the proximity of the three cluster 1, generalizing it to a new Cluster 1 with 3 + 20 = 23 cities.

After generalization, we now narrowed down our selection to 42 – 23 – 4 = 15 cities.

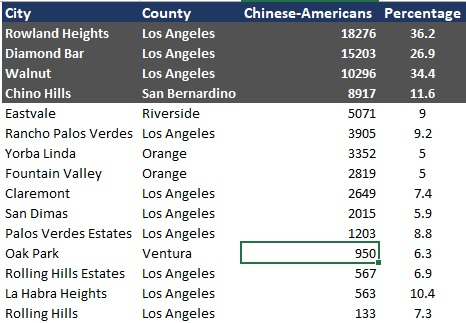
*Table 3: Cluster Generalization of Cities in LA*



**4.5 Top 4 Cities to Launch Business**

Let us go back to the population concentration again, and sort the remaining 16 cities. As shown in the table below, there are 4 cities stand out with Chinese American with about 9,000 or more. We suggest they are the high potential candidates where we shall launch our Chinese restaurant business. They are Rowland Heights, Diamond Bar, Walnut, and Chino Hills.

*Table 4: Tier 1 and Tier 2 Candidates in LA*



**5. Discussion**

Our analysis is conducted at the city level, narrowed down total 88 cities in the Los Angeles area into 4, with additional 11 as next tier hopefuls. To finally start the business, further investigation will be needed and many topics can be resolved by similar technology enlisted in this study. We suggest the future study being focused on the following factors:

* Neighborhood within the city – Zoning, ethical population, crime rate, affluence
* Street location – Parking, accessibility, visibility
* House Facility - Housing price, utilities, equipment
* Niche selection – Upscale or discount business

**Neighborhood**: Zoning is the first fact we need to find out. Some neighborhood simply not allow commercial activities, while other may have strict restrictions that is not desirable for our business. Once we find out the business friendly neighborhood, we shall further understand if there enough local Chinese American ordering take out, if the neighborhood is safe, and if people living in that area having enough disposable cash to dine out.

**Street Location**: Location again, after city location, neighborhood location, and now need to consider exactly which street and which shop to buy. We need to make sure the location has enough parking space to facilitate the customers. The location ideally close to downtown, or have easy access through artery roads. When a potential client standing on the street, best he or she can see the business sign easily.

**House Facility**: This can easily be another ML project to compare the housing price and get a reasonable deal to buy or rent a shop. It is not uncommon that some business failed not because of the business prosperity, but because of skyrocketing rents. Also need to calculate the ROI of utilities, including electricity and water, sewer of the shop. There shall be enough room to install appliance and other equipment, without violating the local codes.

**Niche selection**: Given the recent dramatic demographic changes, one need to decide on going upscale or discount route. As mentioned in the article earlier, there is a strong trend to facilitate fancy and high end Chinese food. Anyhow Chinese food in the mainland, Taiwan, and Hong Kong are very sophisticated, easily in par, if not surpassing, French, Italian, and Greek. It is Chinese Chines food, not American Chinese food, opening up a broad new market. On the other hand, the “old” discount take-out operation is still well received by a big crowd of followers, for its bargain price and convenience.

Of course, there are many other business related issues to consider, such as getting cost effective loan, design a efficient marketing strategy, hire well-know chef and industrial staff, ….beyond the data science arena we discussed.

We expect this study can be further refined with additional information on the Chinese American migration dynamics. Especially, more details in the neighborhood level will help push the study to a level further help the decision making.

**6. Conclusion**

We suggest four cities to launch a new Chinese restaurant, that is, Rowland Heights, Diamond Bar, Walnut, and Chino Hills. There are 11 other cities in Table 4 as tier 2 candidates to consider.

The recommendation is based on the investigation of Chinese American population distribution, as well as the spread of competing business in each cities. We also factored in the proximity of cities, to minimize the competition risks. We believe a city with high concentration of ethical Chinese and less competition will provide the best combination to launch a Chinese restaurant.

**Appendix:**

<https://www.census.gov/>

https://www.migrationpolicy.org/programs/data-hub/us-immigration-trends#source

<https://en.wikipedia.org/wiki/List_of_U.S._cities_with_significant_Chinese-American_populations>

<https://www.cnbc.com/2016/01/20/heres-the-real-reason-why-most-restaurants-fail.html>

<https://www.thebalancesmb.com/choosing-a-location-for-your-restaurant-2888635>

<https://www.latimes.com/food/la-fo-modern-chinese-food-california-20190411-story.html>

<https://nypost.com/2016/01/25/the-17-best-chinese-restaurants-in-america/>

<https://en.wikipedia.org/wiki/List_of_cities_in_Los_Angeles_County,_California>