According to VOA, Chinese food is the most popular foreign food in United States, so it would be a good choice for a Chinese immigrant to open a Chinese restaurant. But the problem is where he or she should open the restaurant? The first reaction is that opening a Chinese restaurant in the city with the most population, but among these big cities, which one is the best choice? This report is trying to solve this question.

Data I will use:

1. List of North American cities by population
2. Geolocation data of each American city acquired from Nominatim
3. Land Area of U.S cities.
4. List of U.S states
5. Number of Chinese Restaurants near a specific geolocation acquired from Foursquare API

Methodology:

The methodology behind this report is based on that how many people can access a Chinese restaurant in a specific area. So to get this,

$$Population\ Density = \frac{City\ Population}{City\ Area}$$

$$Number\ of\ people\ in\ a\ circle\ with\ 3km\ radius\ circle = Population\ Density * Area$$

$$\# = \frac{Number\ of\ people}{Number\ of\ Chinese\ Restaurants}$$

1. First, I need to find top U.S populated city, because I assume to open a new Chinese restaurant in the most populated cities. I get this population data from Wikipedia as I listed above
2. After scraping the data from Wikipedia and cleaned it, I will load the city information to Nominatim package in Python and get the cities' geolocation.
3. I will scrape U.S cities area information from another website.
4. After having city population and city area, I can calculate population density by using city population divided by city area.

5. I will load geolocation data that I got from step 2 to Foursquare API and use the search function to get how many Chinese restaurants there are around 3km circle area.
6. Multiply 3km and population density, we can have the number of people is there in the 3km circle.
7. Now number of people divided by number of Chinese restaurants, we can get the number we want. The larger the number, more people are sharing one Chinese restaurants in that area, which means it is a good choice to open a new Chinese restaurant in that area. On the contrary, the smaller the number, less people are sharing a Chinese restaurant, and it is not a good choice to open a new Chinese restaurant.
8. Mapping the result to map. As mentioned above, the bigger the circle, the better choice to open a restaurant, the smaller the circle, the worse to open a restaurant.

Results:

Surprisingly, Portland, Oregon has the least number, which means there are most Chinese restaurants. Los Angles places the third because there are many Chinese groups living in that area and it is not surprising that there are lots of Chinese restaurants.

Recommendation:

I would recommend opening a new Chinese restaurant in Columbus OH, Fresno CA and Jacksonville FL because these cities has the highest number and it means more people are sharing only one Chinese restaurants. Places with small circles such as Portland, Seattle and Los Angles should be avoided because there are more Chinese restaurants and the competition would be tough.

Conclusion:

This is a simple analysis and I think the results could be better if I have more data. For example, Foursquare only gives 50 venues for free accounts, and places such as New York City may have more than 50 Chinese Restaurants within 3km area. New York City in this analysis places in the last 5, which is absurd because there are lots of Chinese restaurants in the city.