**Distribution of restaurants in Los Angeles**

Kangwoo Lee

3.16.2021

**Introduction**

1. Background

Los Angeles is known as a multicultural city with many ethnic enclaves, and a city with many gourmet restaurants. Los Angeles hosts huge populations of Cambodians, Persians, Armenians, Belizeans, Bulgarians, Ethiopians, Filipinos, Guatemalans, Hungarians, Koreans, Israelis, Mexicans, Salvadorans, Thais, Japanese, Chinese and Pacific Islanders such as Samoans. There are countless number of good restaurants from different cultural background in Los Angeles, and many restaurants have been opened and closed. The diversity of cultural background in a city could be an advantage for a new vendor. Target customers are opened to multicultural choices as they have been adapting to the society. However, it is difficult to choose a cutting-edge to be a long-lasting restaurant in this environment as there are numerous choices for the decision making. Therefore, there should be a distinguished strength for a new restaurant to start their business.

1. Problem

How can we define where and what kind of restaurant should I open a restaurant? We start from understanding the distribution of each type of restaurant in Los Angeles, and map the locations of restaurants to solve this question. There are numerous restaurants in Los Angeles, and it is impossible to list all restaurant in a list manually.

1. Interest

To open a restaurant, we need to check: 1) population, 2) ethnic enclaves, 3) information of venues in the area. Those blow-in vendors who wants to open a restaurant and lacks information of venues in Los Angeles would have better ideas with this program.

**Method**

1. Data acquisition

All the information about venues are automatically extracted from Foursquare API.

2. Data processing

The information of venues were extracted from Foursquare API. The coordinate of Los Angeles, the radius of searching area are included into a searching query, and it is passed to Foursquare API to extract the data.

Data are classified by categories of each venue. Venues without clear categorical information are excluded in this project.

1. Feature selection

We classify each venue by its category, then we evaluate the frequency of each category type. There are many different kinds of category names in the data, and we exclude venues with random category names or unrelated names. (e.g. ‘Construction & Landscaping’, ‘None’)

Furthermore, our primary interest is to avoid the location in which many similar restaurants are open already. We extract categories with multiple venues from the raw data, to see which area has multiple restaurants with similar concepts.

**Result**

|  |  |
| --- | --- |
| **Category** | **Frequency** |
| French restaurant | 3 |
| American restaurant | 2 |
| New American restaurant | 2 |
| Mexican restaurant | 6 |
| Chinese restaurant | 5 |
| Italian restaurant | 4 |
| Asian restaurant | 2 |
| Latin American restaurant | 2 |
| Mediterranean restaurant | 2 |
| Japanese retaurant | 4 |

Table 1. The frequency of each venue type

There are 10 most frequent categories of venues: French Restaurant, American Restaurant, New American Restaurant, Mexican Restaurant, Chinese Restaurant, Italian Restaurant, Asian Restaurant, Latin American Restaurant, Mediterranean Restaurant, and Japanese Restaurant (Table 1). Restaurants without information uploaded on Foursquare API are not included.

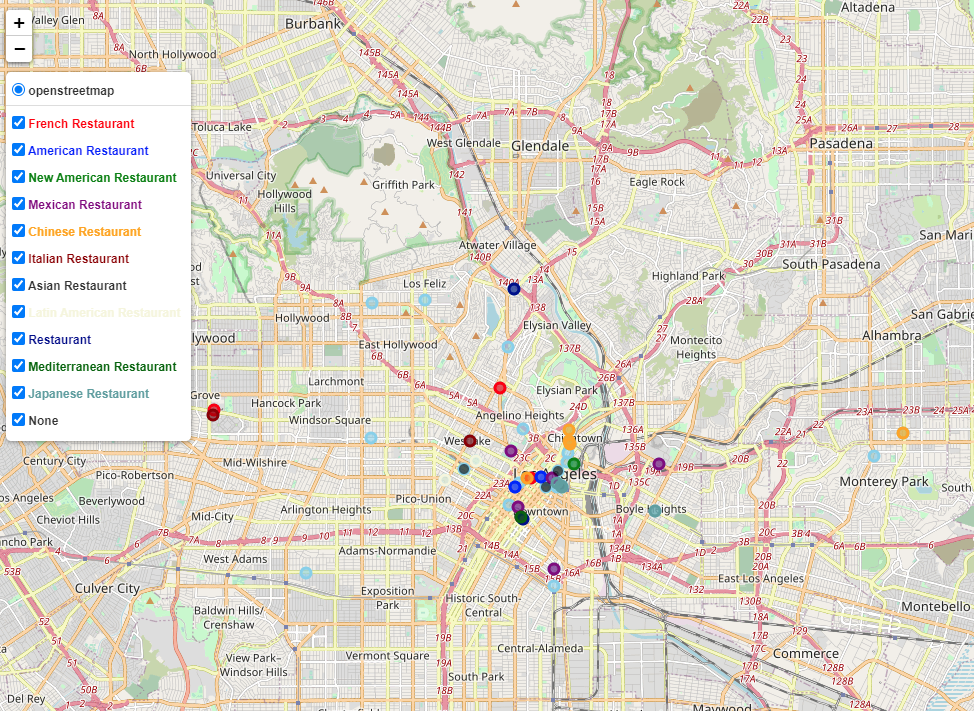


Figure 1. Restaurant locations. Each color indicates venue types. Skyblue circles indicates labeled as ‘none’ in the category information.

**Discussion**

In Figure 1, we can see the distribution of restaurants in Los Angeles. There are many restaurants around Los Angeles downtown, and they compete with other venues by its distinguished characteristics. We can observe that similar types of restaurants would not open in a close distance.

**Conclusion**

We can observe the distribution of each restaurant by an auto-extraction process from Foursquare API data. There are a few limitations which needs to be further improved.

First of all, many venues lack information on their keywords. For example, there are many more restaurants in Los Angeles, but not all restaurants are listed in the map nor dataset. More detailed pre-processing is required.

Second, category information needs to be updated. Many venues have an empty category, or insufficient information are presented, such as ‘Restaurant’, ‘None’. These two issues should be further analyzed through more steps of pre-processing.