

Capstone Project - The Battle of Neighborhoods

1. Introduction

Taipei is the capital and a special municipality of Taiwan (officially the Republic of China, ROC). It is also a city whose allure lies in its blend of Chinese culture with a curious fusion of Japanese, Southeast Asian and American influences. Whenever visitors visit to a new city they are looking for the restaurants during their stay. Thus, in this presentation, I will explore restaurants of Taipei based on Foursquare data. Try to provide support to new visitors helping them to explore the places visiting Taipei could explore.

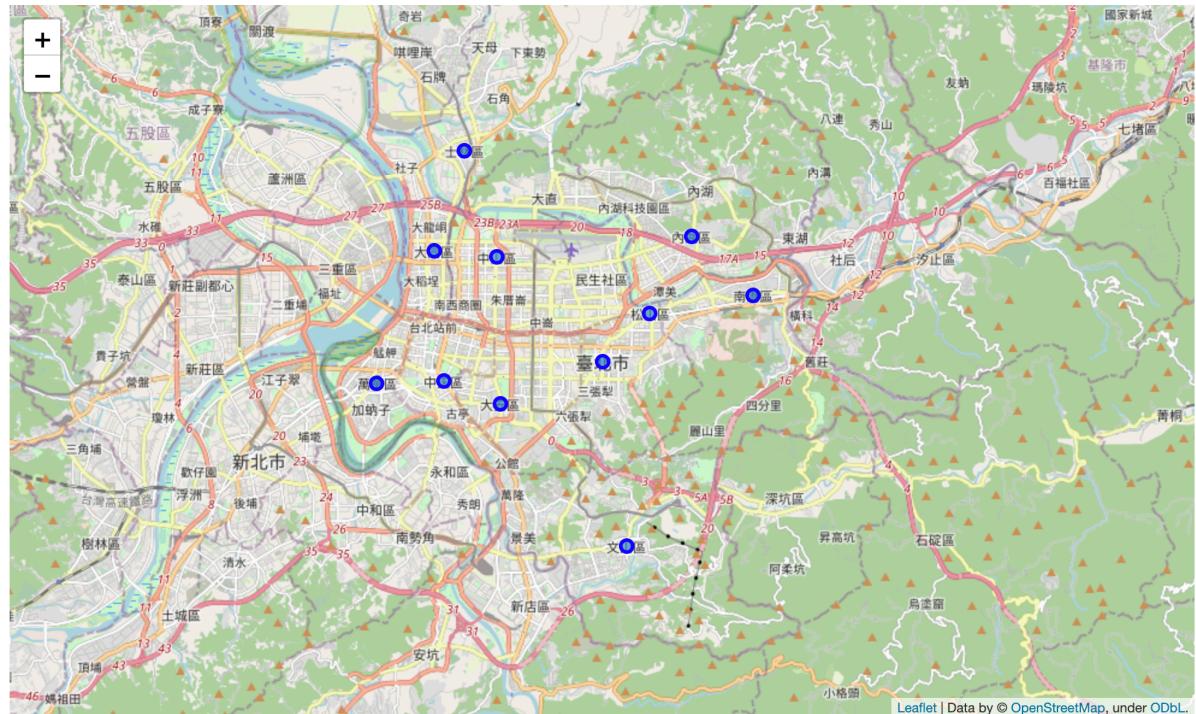
2. Data Requirements

- ✓ Taipei Administrative Divisions
<https://en.wikipedia.org/wiki/Taipei>
- ✓ Restaurants in each districts of Taipei
Use Foursquare APIs to get restaurants(section=food) in each district.

The first step was to get Taipei districts table from Wikipedia and transformed it to pandas dataframe. Using geocoder class of Geopy client to get coordinates of those districts. Then, I used Foursquare API to get the top 100 restaurants of all districts within a radius of 5000 meters. As the result, totally got 1160 restaurants.

	Name	Chinese	Population	Area	Postalcode	Latitude	Longitude
0	Beitou	台北市北投區	257922	56.8216	112	25.132419	121.501379
1	Daan (Da-an, Da'an)	台北市大安區	312909	11.3614	106	25.026515	121.534395
2	Datong	台北市大同區	131029	5.6815	103	25.065986	121.515514
3	Nangang (Nankang)	台北市南港區	122296	21.8424	115	25.054578	121.606600
4	Neihu	台北市內湖區	287726	31.5787	114	25.069664	121.588998
5	Shilin	台北市士林區	290682	62.3682	111	25.091840	121.524207
6	Songshan	台北市松山區	209689	9.2878	105	25.049885	121.577272
7	Wanhua	台北市萬華區	194314	8.8522	108	25.031933	121.499332
8	Wenshan	台北市文山區	275433	31.5090	116	24.989786	121.570458
9	Xinyi	台北市信義區	229139	11.2077	110	25.037520	121.563680
10	Zhongshan	台北市中山區	231286	13.6821	104	25.064361	121.533468
11	Zhongzheng	台北市中正區	162549	7.6071	100	25.032361	121.518267

List of 12 Taipei Districts



Visualize geographic details of 12 Taipei Districts

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Beitou	25.132419	121.501379	金春發牛肉麵	25.137720	121.493171	Asian Restaurant
1	Beitou	25.132419	121.501379	拾米屋 SheMe House	25.136224	121.499005	Café
2	Beitou	25.132419	121.501379	張吳記什錦麵	25.127777	121.500023	Noodle House
3	Beitou	25.132419	121.501379	阿二麻辣旗艦店	25.116681	121.516192	Hotpot Restaurant
4	Beitou	25.132419	121.501379	東方泰國小館	25.114670	121.515385	Thai Restaurant
...
1155	Zhongzheng	25.032361	121.518267	Robin's Grill	25.054178	121.524381	Steakhouse
1156	Zhongzheng	25.032361	121.518267	聞山咖啡	25.020409	121.533680	Café
1157	Zhongzheng	25.032361	121.518267	Yong He Soy Milk King (永和豆漿大王)	25.029858	121.543424	Chinese Breakfast Place
1158	Zhongzheng	25.032361	121.518267	Orange Shabu Shabu House (橘色涮涮屋)	25.038602	121.546190	Shabu-Shabu Restaurant
1159	Zhongzheng	25.032361	121.518267	Kanokwan (老麵攤)	25.052861	121.528862	Thai Restaurant

1160 rows x 7 columns

Sample of restaurants got from Foursquare API

3. Methodology

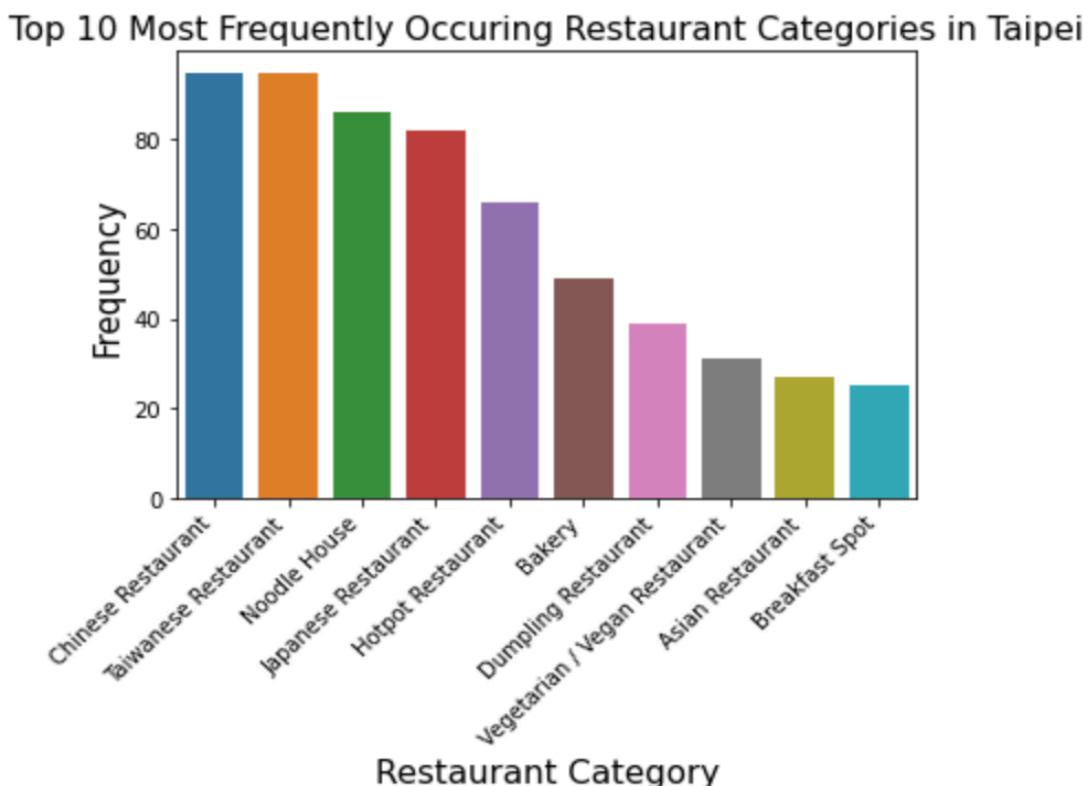
As the first step, I retrieved the data from Foursquare API by extracting venue information from 12 districts of Taipei, upto a distance of 5 Km. After that, I got 1160 restaurants with different categories.

I analyzed the data based on the category of each restaurants. I identified the top category types and places where many restaurants are located so that any visitor can go to one place and enjoy the option to choose amongst many restaurant options. After that, I clustered the districts based on the available information of each restaurants. This will allow us to clearly identify which districts can be recommended and with what characteristics.

3.1.1 Exploratory Data Analysis

Category

We can find out 68 unique categories of restaurants in Taipei. Chinese Restaurants are the top one, the second one is Taiwanese Restaurants, and the third one is Noodle House as we can see in the plot below:

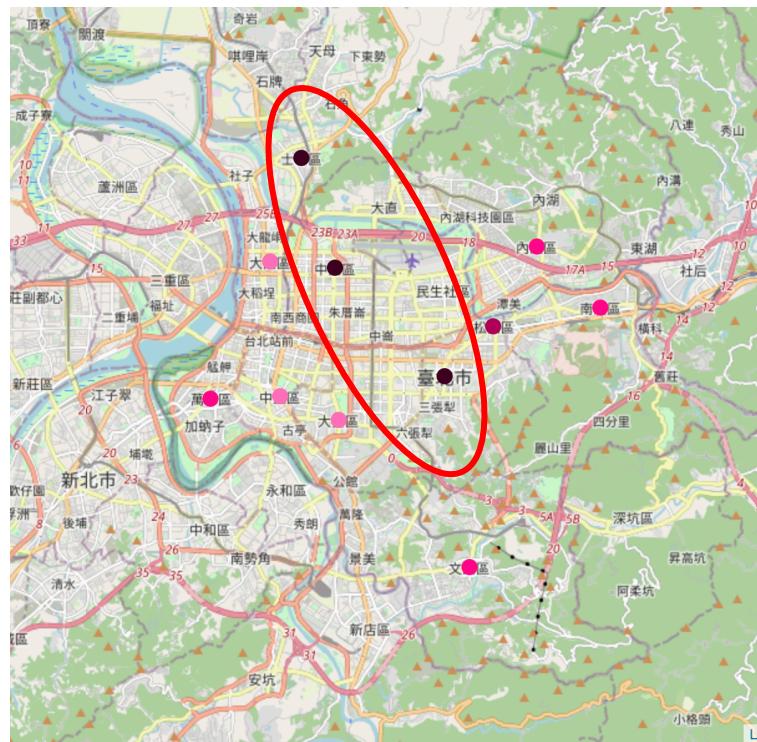


As the result, when visitor visit Taipei definitely need to try the Chinese Restaurants and Taiwanese Restaurants.

Category vs. Location

We can find out each district has about 30 different categories. Songshan, Shillin, and Zhongshan are the top 3 districts with many different categories.

	Name	Category Count
6	Songshan	39
5	Shilin	38
10	Zhongshan	38
4	Neihu	37
3	Nangang (Nankang)	35
7	Wanhua	35
8	Wenshan	35
9	Xinyi	34
1	Daan (Da-an, Da'an)	33
2	Datong	33
11	Zhongzheng	33
0	Beitou	27



Output each neighborhood along with the top 5 most common categories:

----Beitou----

	venue	freq
0	Chinese Restaurant	0.19
1	Fast Food Restaurant	0.08
2	Japanese Restaurant	0.06
3	Italian Restaurant	0.06
4	American Restaurant	0.04

----Datong----

	venue	freq
0	Chinese Restaurant	0.14
1	Taiwanese Restaurant	0.11
2	Noodle House	0.11
3	Japanese Restaurant	0.08
4	Hotpot Restaurant	0.08

----Zhongshan----

	venue	freq
0	Japanese Restaurant	0.12
1	Chinese Restaurant	0.10
2	Taiwanese Restaurant	0.10
3	Hotpot Restaurant	0.07
4	Asian Restaurant	0.06

----Daan (Da-an, Da'an)----

	venue	freq
0	Chinese Restaurant	0.13
1	Japanese Restaurant	0.09
2	Bakery	0.09
3	Taiwanese Restaurant	0.08
4	Noodle House	0.06

----Nangang (Nankang)----

	venue	freq
0	Chinese Restaurant	0.11
1	Noodle House	0.11
2	Taiwanese Restaurant	0.09
3	Japanese Restaurant	0.08
4	Hotpot Restaurant	0.06

----Zhongzheng----

	venue	freq
0	Taiwanese Restaurant	0.12
1	Noodle House	0.11
2	Chinese Restaurant	0.09
3	Bakery	0.07
4	Japanese Restaurant	0.06

----Neihu----

	venue	freq
0	Noodle House	0.09
1	Taiwanese Restaurant	0.08
2	Chinese Restaurant	0.08
3	Japanese Restaurant	0.08
4	Hotpot Restaurant	0.06

----Songshan----

	venue	freq
0	Noodle House	0.11
1	Hotpot Restaurant	0.09
2	Japanese Restaurant	0.08
3	Taiwanese Restaurant	0.08
4	BBQ Joint	0.05

----Wenshan----

	venue	freq
0	Chinese Restaurant	0.11
1	Japanese Restaurant	0.09
2	Noodle House	0.09
3	Taiwanese Restaurant	0.07
4	Dumpling Restaurant	0.06

----Shilin----

	venue	freq
0	Chinese Restaurant	0.14
1	Hotpot Restaurant	0.09
2	Taiwanese Restaurant	0.09
3	Japanese Restaurant	0.08
4	Thai Restaurant	0.05

----Wanhua----

	venue	freq
0	Taiwanese Restaurant	0.15
1	Noodle House	0.13
2	Bakery	0.08
3	Japanese Restaurant	0.06
4	Chinese Restaurant	0.06

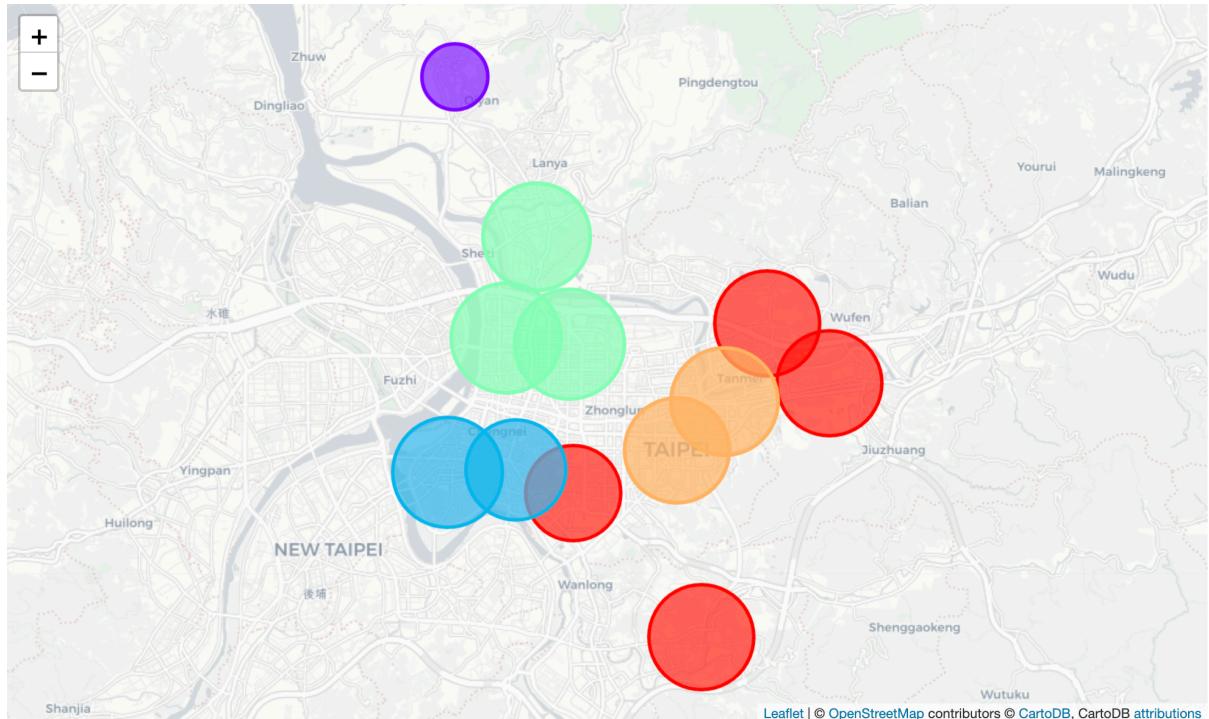
----Xinyi----

	venue	freq
0	Taiwanese Restaurant	0.10
1	Noodle House	0.09
2	Japanese Restaurant	0.09
3	Hotpot Restaurant	0.09
4	Bakery	0.07

Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Beitou	Chinese Restaurant	Fast Food Restaurant	Japanese Restaurant	Italian Restaurant	Restaurant	Bakery	Dumpling Restaurant	Hotpot Restaurant	Korean Restaurant
1	Daan (Da-an, Da'an)	Chinese Restaurant	Japanese Restaurant	Bakery	Taiwanese Restaurant	Hotpot Restaurant	Noodle House	Breakfast Spot	Vegetarian / Vegan Restaurant	Dumpling Restaurant
2	Datong	Chinese Restaurant	Noodle House	Taiwanese Restaurant	Hotpot Restaurant	Japanese Restaurant	Asian Restaurant	Bakery	Sushi Restaurant	Ramen Restaurant
3	Nangang (Nankang)	Noodle House	Chinese Restaurant	Taiwanese Restaurant	Japanese Restaurant	Hotpot Restaurant	Dumpling Restaurant	Bakery	Sushi Restaurant	Vietnamese Restaurant
4	Neihu	Noodle House	Japanese Restaurant	Chinese Restaurant	Taiwanese Restaurant	Hotpot Restaurant	Bakery	Sushi Restaurant	Dumpling Restaurant	French Restaurant
5	Shilin	Chinese Restaurant	Taiwanese Restaurant	Hotpot Restaurant	Japanese Restaurant	Thai Restaurant	Asian Restaurant	Dumpling Restaurant	Restaurant	Breakfast Spot
6	Songshan	Noodle House	Hotpot Restaurant	Japanese Restaurant	Taiwanese Restaurant	Vegetarian / Vegan Restaurant	BBQ Joint	Sushi Restaurant	Seafood Restaurant	Bakery
7	Wanhua	Taiwanese Restaurant	Noodle House	Bakery	Chinese Restaurant	Japanese Restaurant	Hotpot Restaurant	Food Truck	Asian Restaurant	Ramen Restaurant
8	Wenshan	Chinese Restaurant	Japanese Restaurant	Noodle House	Taiwanese Restaurant	Dumpling Restaurant	Hotpot Restaurant	Fast Food Restaurant	Italian Restaurant	Breakfast Spot
9	Xinyi	Taiwanese Restaurant	Japanese Restaurant	Hotpot Restaurant	Noodle House	Bakery	Vegetarian / Vegan Restaurant	Dumpling Restaurant	BBQ Joint	Breakfast Spot
10	Zhongshan	Japanese Restaurant	Chinese Restaurant	Taiwanese Restaurant	Hotpot Restaurant	Asian Restaurant	Noodle House	Seafood Restaurant	Dumpling Restaurant	Ramen Restaurant
11	Zhongzheng	Taiwanese Restaurant	Noodle House	Chinese Restaurant	Bakery	Japanese Restaurant	Hotpot Restaurant	Dumpling Restaurant	Food Truck	Snack Place
										Hakka Restaurant

3.1.2 Clustering Analysis

Finally, I tried to cluster these 12 districts based on the venue categories and use K-Means clustering. So, based on the similarities of venue categories, these 12 districts will be clustered into 5 groups. The 5 clusters in a map as below:



4. Results & Discussion

I found out some interesting insights which might be useful to visitors. Let's summarize the findings:

- Chinese restaurants are the most popular restaurants in the 12 districts.
- Songshan has maximum number categories of restaurants.
- The districts in the central of the city has more different categories of restaurant than in the border of the city.
- The clustering was based on the category of restaurants on each district, so districts in the same cluster indicate that each of those districts presents a similar experience to the visitor in terms of category of food. Visitor can plan to visit different clusters to experience different food.

In the analysis, I ignored many other factors like prices of restaurants, rating of restaurant, distance of the restaurants from stations, and so on. Hence, this analysis only helps visitors to get an overview of restaurants distribution by categories. This results also could potentially vary if we use some other clustering techniques like DBSCAN.

5. Conclusion

The purpose of this project was to explore the districts of Taipei that can help visitor visiting Taipei to fit the most their needs. I have made to parse web content to get districts of Taipei, use Foursquare API to explore restaurants in each districts and saw the results of segmentation of districts.