

## Problem:

The location is very important when you want to open a restaurant in Hong Kong. Traditionally people would look for traffic pattern information, demographic and lifestyle data online or by handing out survey. There are real-time data sources for traffic and demographic information. I would use the Data Science technique like Clustering and Visualization to solve the problem.

## Data:

In the first part of the project, I draw data frames containing neighborhood names and their locations. In the second part of the project, I use the Four Square API to explore more information. Foursquare is a technology company that built a massive dataset of location.

## Methodology:

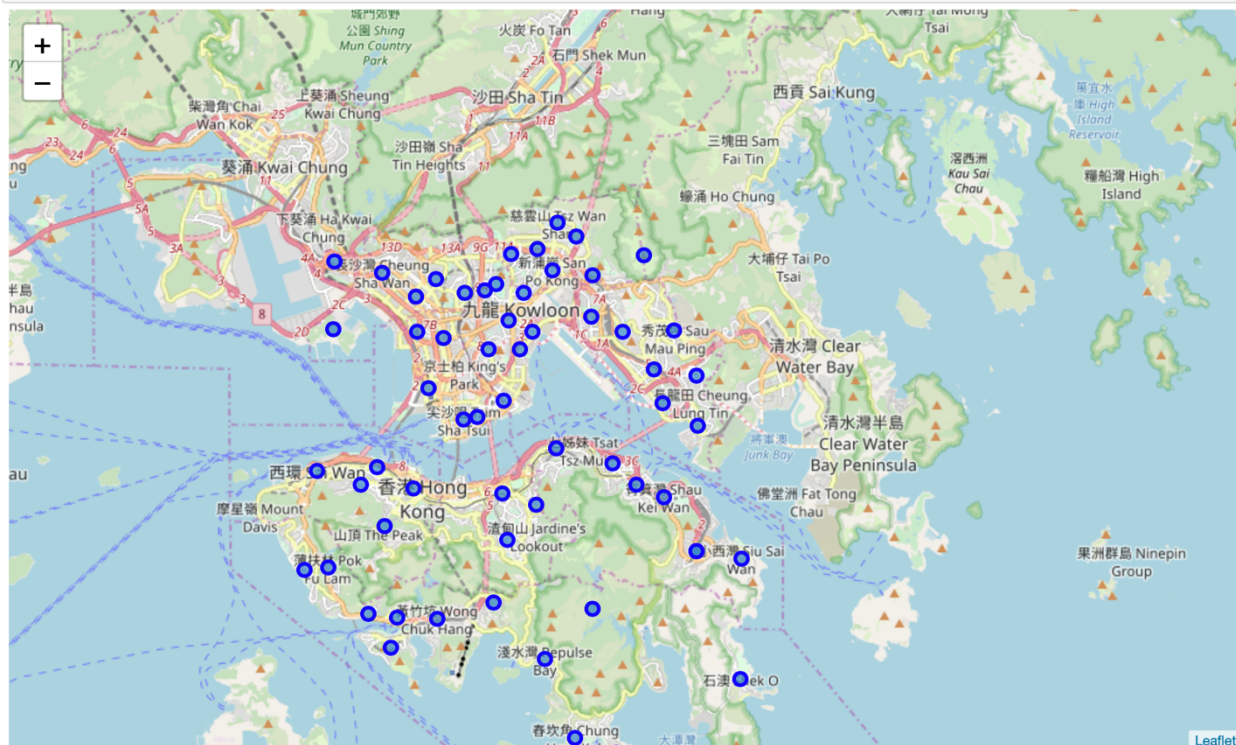
The ultimate goal of this project is to provide an easy-to-understand report to those looking for the right location to open a restaurant in Hong Kong. The first step is to collect geospatial data of the city. Hong Kong consists of Hong Kong Island, the Kowloon Peninsula, the New Territories, Lantau Island, and over 200 other islands. This project will focus on Hong Kong Island and Kowloon. Lucky enough, we can find the csv file of detailed neighborhood information of Hongkong. We directly import it as our main data frame and begin to deal with it.

```
In [3]: df_hk = pd.read_csv('neighborhoods_hong_kong.csv')
df_hk
```

Out[3]:

	District	Neighborhood	Latitude	Longitude
0	Central & Western	Central District	22.281322	114.160258
1	Central & Western	Mid-Levels	22.282405	114.145809
2	Central & Western	The Peak	22.272003	114.152417
3	Central & Western	Sai Wan	22.285838	114.134023
4	Central & Western	Sheung Wan	22.286870	114.150267
5	Eastern	Chai Wan	22.265607	114.237964
6	Eastern	North Point	22.291657	114.199545
7	Eastern	Quarry Bay	22.287755	114.214932
8	Eastern	Sai Wan Ho	22.282446	114.221506
9	Eastern	Shau Kei Wan	22.279343	114.228898
10	Eastern	Siu Sai Wan	22.263727	114.250208
11	Southern	Aberdeen	22.248769	114.155954
12	Southern	Ap Lei Chau	22.241210	114.154136
13	Southern	Chung Hom Kok	22.218262	114.204760
14	Southern	Cyberport	22.260697	114.130286
15	Southern	Deep Water Bay	22.252438	114.182113
16	Southern	Pok Fu Lam	22.261357	114.136809
17	Southern	Tin Wan	22.249639	114.148008
18	Southern	Repulse Bay	22.238108	114.196318
19	Southern	Stanley	22.214391	114.215913
20	Southern	Shek O	22.233167	114.250033
21	Southern	Tai Tam	22.250999	114.209394

After collecting and cleaning the data, I plot the neighborhoods on a map.



Now I begin exploring the Four Square API. I want to get the top 100 venues that are in a certain neighborhood within a radius of 1000 meters. I'll repeat the process for all the neighborhoods. This radius can be changed to meet different requirements. However if you want to extend the 100 venues to more venues you will need to purchase the professional version of Four Square API.

Here are the top five venues in Central District

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Central District	22.281322	114.160258	Mandarin Oriental Hong Kong (香港文華東方酒店)	22.281879	114.159443	Hotel
1	Central District	22.281322	114.160258	Mandarin Grill + Bar (文華扒房 + 酒吧)	22.281462	114.160156	Steakhouse
2	Central District	22.281322	114.160258	Mott 32 (卅二公館)	22.280696	114.159380	Dim Sum Restaurant
3	Central District	22.281322	114.160258	Dr. Fern's Gin Parlour	22.280985	114.158391	Lounge
4	Central District	22.281322	114.160258	The Mandarin Cake Shop	22.281959	114.159416	Bakery

Display the top 10 most common venues for each neighborhood.

	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	Aberdeen	Chinese Restaurant	Cha Chaan Teng	Athletics & Sports	Coffee Shop	Thai Restaurant	Market	Grocery Store	Food Court	Park	Fast Food Restaurant
1	Ap Lei Chau	Chinese Restaurant	Furniture / Home Store	Fast Food Restaurant	Shopping Mall	Market	Clothing Store	Pet Store	Park	Café	Paper / Office Supplies Store
2	Causeway Bay	Japanese Restaurant	Sushi Restaurant	Dessert Shop	Coffee Shop	Bakery	Chinese Restaurant	Hotel	Bubble Tea Shop	Sporting Goods Shop	Bookstore
3	Central District	Chinese Restaurant	Italian Restaurant	Steakhouse	Coffee Shop	Social Club	Gym / Fitness Center	Hotel	Hotel Bar	Cantonese Restaurant	Lounge
4	Cha Kwo Ling	Noodle House	Convenience Store	Fast Food Restaurant	Shanghai Restaurant	Soccer Field	Dim Sum Restaurant	Department Store	Fruit & Vegetable Store	Frozen Yogurt Shop	Fried Chicken Joint

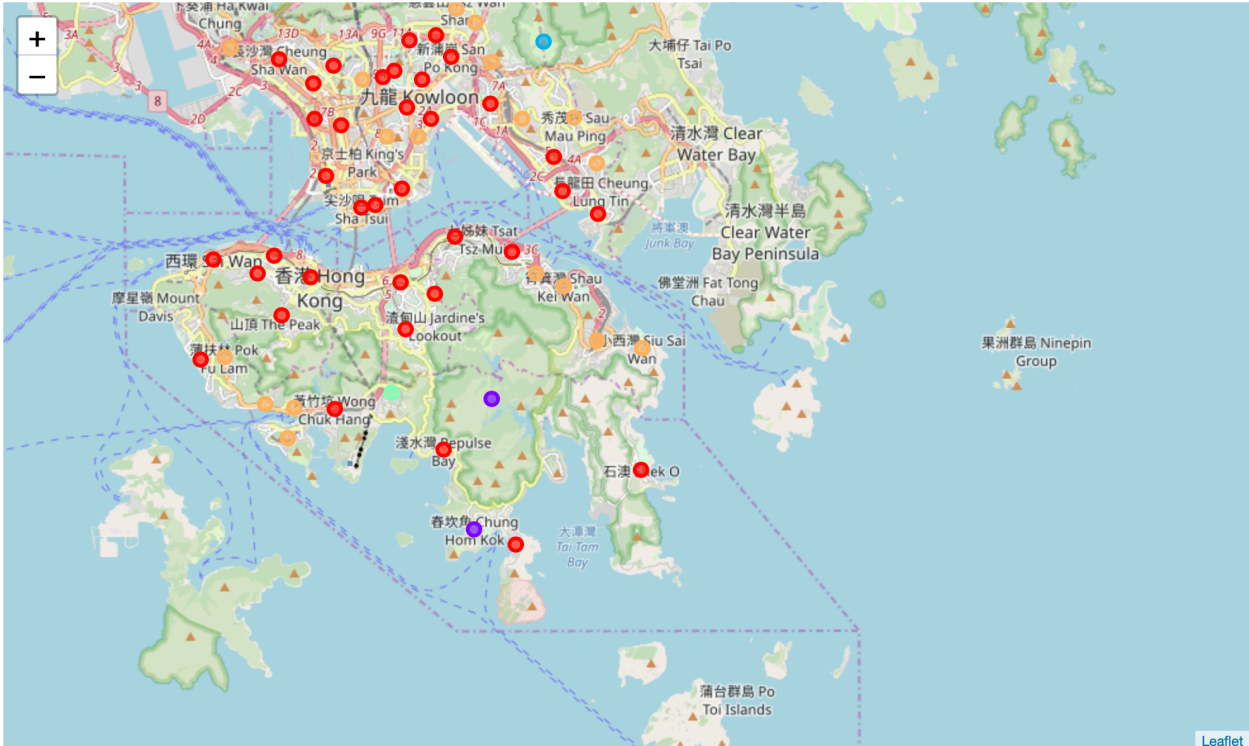
Now we have much useful information. If I want to open a Chinese restaurant I might consider Aberdeen or Ap Lei Chau because Chinese restaurants are popular in these two neighborhoods. If I want to open a sushi restaurant I might consider Causeway for the same reason. If the number of neighborhoods is very large it might be a good idea to cluster the neighborhoods.

We also know that traffic is also a crucial factor in deciding where to open the restaurant, so we append the traffic information to our data frame.

	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	Station
0	Aberdeen	Chinese Restaurant	Cha Chaan Teng	Athletics & Sports	Coffee Shop	Thai Restaurant	Market	Grocery Store	Food Court	Park	Fast Food Restaurant	Yes
1	Ap Lei Chau	Chinese Restaurant	Furniture / Home Store	Fast Food Restaurant	Shopping Mall	Market	Clothing Store	Pet Store	Park	Café	Paper / Office Supplies Store	Yes
2	Causeway Bay	Japanese Restaurant	Sushi Restaurant	Dessert Shop	Coffee Shop	Bakery	Chinese Restaurant	Hotel	Bubble Tea Shop	Sporting Goods Shop	Bookstore	No
3	Central District	Chinese Restaurant	Italian Restaurant	Steakhouse	Coffee Shop	Social Club	Gym / Fitness Center	Hotel	Hotel Bar	Cantonese Restaurant	Lounge	No
4	Cha Kwo Ling	Noodle House	Convenience Store	Fast Food Restaurant	Shanghai Restaurant	Soccer Field	Dim Sum Restaurant	Department Store	Fruit & Vegetable Store	Frozen Yogurt Shop	Fried Chicken Joint	No

Finally, I use k-means to cluster the neighborhoods into 5 clusters. They are represented using different colors. If you click on the small dot a label will pop up which tells you the name and

number of the cluster:



For each cluster I draw a data frame:

Cluster 1

```
hk_merged.loc[hk_merged['Cluster Labels'] == 0, hk_merged.columns[[1] + list(range(5, hk_merged.shape[1]))]]
```

	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	Station
0	Central District	Chinese Restaurant	Italian Restaurant	Steakhouse	Coffee Shop	Social Club	Gym / Fitness Center	Hotel	Hotel Bar	Cantonese Restaurant	Lounge	No
1	Mid-Levels	Thai Restaurant	Café	Noodle House	Italian Restaurant	Japanese Restaurant	Park	Coffee Shop	Dessert Shop	Seafood Restaurant	Steakhouse	No
2	The Peak	Scenic Lookout	Ice Cream Shop	Shopping Mall	Asian Restaurant	Restaurant	Supermarket	Sushi Restaurant	Gift Shop	Grocery Store	Pizza Place	Yes
3	Sai Wan	Dessert Shop	Noodle House	Pier	Malay Restaurant	Spanish Restaurant	New American Restaurant	Multicuisine Indian Restaurant	Café	Boxing Gym	Furniture / Home Store	No
4	Sheung Wan	Café	Japanese Restaurant	Italian Restaurant	Coffee Shop	Restaurant	Chinese Restaurant	Bar	Supermarket	Indian Restaurant	Food Court	No
6	North Point	Burger Joint	Thai Restaurant	Noodle House	Hong Kong Restaurant	Japanese Restaurant	Hotpot Restaurant	Café	Dim Sum Restaurant	Park	Gastropub	No
7	Quarry Bay	Café	Coffee Shop	Japanese Restaurant	Department Store	Thai Restaurant	Chinese Restaurant	Vietnamese Restaurant	Ice Cream Shop	Food Court	Taiwanese Restaurant	No
14	Cyberport	Coffee Shop	Bus Stop	Gym	Hotel Bar	Sports Club	Business Service	Café	Supermarket	Multiplex	Cantonese Restaurant	Yes
18	Repulse Bay	Café	Pizza Place	Chinese Restaurant	Supermarket	Seafood Restaurant	Spa	Bus Stop	Bus Station	Gastropub	Shopping Mall	Yes
19	Stanley	Surf Spot	Beach	History Museum	Playground	Fast Food Restaurant	Fruit & Vegetable Store	Frozen Yogurt Shop	Fried Chicken Joint	French Restaurant	Food Court	No
20	Shek O	Thai Restaurant	BBQ Joint	Café	Beach	Noodle House	Breakfast Spot	Golf Course	French Restaurant	Bar	Mediterranean Restaurant	Yes

**Conclusion:**

Look back to our visualization map, we can exclude Cluster 3,4,5 from ur candidates since there are mountains and parks. After examining cluster 1 and cluster 2, I'd like to say cluster 1 represents residential area and cluster 2 represents commercial area. So if you want to open a restaurant, you should firstly choose the type of your restaurant and choose the corresponding location.