The Battle of Neighbourhoods

Discover Central Singapore: Where to open a bubble tea shop?

By: Ke Guanguan

May 2020

Table of Content

1. Introduction

Background

Nowadays, more and more young people, as well as many other various groups of citizens, are indulging themselves in bubble tea. Bubble tea has become so popular a drink that may surpass other beverages like coffee or tea. In central Singapore, there are many bubble tea shops. For instance, Yi Dian Dian is a renowned local brand, and Hey Tea is a Chinese brand which gains its fame swiftly in the past few months.

More and more bubble tea stores are being constructed in Singapore. However, there are many problems which should be considered when opening one. To determine the location is a vital part of running a bubble tea shop.

Business problems

This project is designed to tackle with two business problems.

Firstly, it helps investors to choose the suitable location in central Singapore to start a bubble tea shop where they can gain more profit.

Secondly, it further helps citizens to choose the residential area by giving one more criterion: are there any bubble tea shops nearby my residence? As for business, it also gives light to real estate companies. For example, they may choose to build more facilities catering for young people where there are some bubble tea shops nearby.

Targeted audience

As mentioned above, the project mainly serves two business problems, which concerns potential bubble tea runners, consumers and real estate companies.

Investors may use this to determine where to build the bubble tea shop in central Singapore according to the density of bubble tea shops around one particular PA. For citizens, they may use this to check whether there are bubble shops nearby the neighbourhoods. For real estate companies, they can utilize the growing trend of bubble tea consumption and opening of shops to decide which kind of facilities to build within the PA.

2. Data

Following data is needed to develop the projects:

- List of PAs in central Singapore: In this project, we specifically investigate central Singapore.
- The geographic coordinates of each PA: Geographic information of each PA can help to get the venue data.
- Venue data: Information help us to cluster the PA and thus determine the location of opening a bubble tea shop.

Data sources:

We get the list of PAs in central Singapore from Wikipedia (https://en.wikipedia.org/wiki/List_of_places_in_Singapore). We use Geocoder to gain the geographic information of the PA. Then Foursquare API is needed to discover the venue information of respective places.

We use the K-means method to cluster the PAs using machine learning method, and view the map use Folium.

3. Methodology

At first, we should scrap list of PAs in central Singapore from Wikipedia. We create the data frame which contains the name of PAs.

To further investigate the venues, we need geographical coordinates of respective PA. We gain the latitude and longitude using Geocoder, and merge the list of those data with the list of PAs. Thus, we get the fundamental information to investigate the nearby bubble tea shops.

We use Folium to visualize the PA on the map to check whether the geographical coordinates are located correctly.

Then we request information from Foursquare API to get the 100 venues within the radius of 5 kilometres. Venue name, venue category, and geographic coordinate have been gained after the API calls. We arrange the data within the unique category and counts the frequency of each category.

Thus, we get the counts of bubble tea shops near each PA. After determining the PA name and count, we can cluster those PA using K-means method. The K chosen in our project is 3 according to the frequency of bubble tea shops appearing in each PA. This clusters the PAs into 3 groups, which helps us to decide which place is the densest of bubble tea shops and which place is not.

We visualize the clusters on the map, which gives the investors and the citizens more intuitive impression of the distribution of bubble tea shops. They can choose the most suitable place to open the bubble tea shop.

4. Results

The K-means clustering shows 3 groups of PAs based on the frequency of the bubble tea shops.

Cluster 0 (red) is PAs with moderate number of bubble tea shops.

	PA	Bubble Tea Shop	Cluster Labels	Latitude	Longitude
0	Bishan	0.010000	0	1.350790	103.851100
17	Singapore River	0.010000	0	1.289710	103.849640
16	Rochor	0.010000	0	1.304130	103.850290
15	River Valley	0.010000	0	1.296855	103.834348
13	Outram	0.010000	0	1.289241	103.835002
12	Orchard	0.010000	0	1.301090	103.839650
11	Novena	0.010000	0	1.319100	103.843720
20	Tanglin	0.010000	0	1.299751	103.817357
10	Newton	0.010000	0	1.312180	103.839120
7	Marina Parade	0.010000	0	1.302515	103.904342
6	Marina East	0.010000	0	1.295790	103.895440
5	Kallang	0.010000	0	1.309415	103.866730
4	Geylang	0.010000	0	1.311470	103.882180
3	Downtown Core	0.012987	0	1.377160	103.955530
8	Marina South	0.010000	0	1.278570	103.857620

Cluster 1 (Purple) is Pas with no or low number of bubble tea shops.

	PA	Bubble Tea Shop	Cluster Labels	Latitude	Longitude
9	Museum	0.0	1	1.301160	103.771950
14	Queenstown	0.0	1	1.299660	103.801720
2	Bukit Timah	0.0	1	1.340410	103.772210
1	Bukit Merah	0.0	1	1.283220	103.816760
18	Southern Islands	0.0	1	1.366670	103.800000
19	Straits View	0.0	1	1.279863	103.853595

Cluster 2 (green) is PA with considerate numbers of bubble tea shops, which contains only one PA. That is **Toa Payoh**.

	PA	Bubble Tea Shop	Cluster Labels	Latitude	Longitude
21	Toa Payoh	0.02	2	1.33448	103.85108

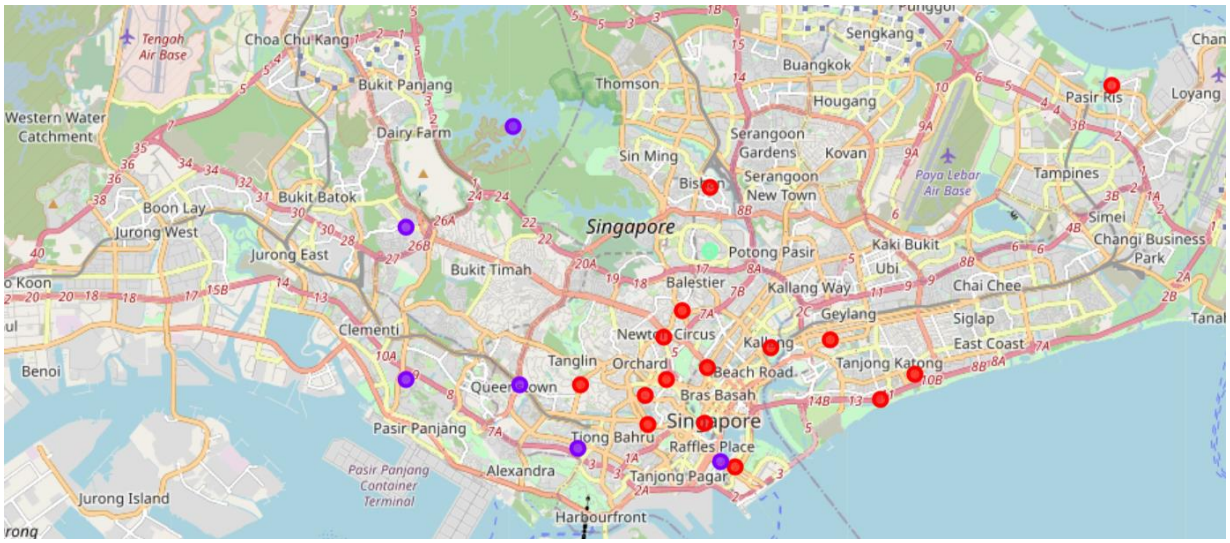

Figure 1 **Central Singapore with clusters**

5. Discussion

As shown in the map, it is obvious that Toa Payoh (**cluster 2**) has the densest bubble tea shops, which is definitely not suitable for those investors to run new bubble tea shops. As the competition between different bubble tea shops increase, it is hard to gain more profit within the area.

However, for those bubble tea lovers, Toa Payoh is a good place to be their residence, as they can access to bubble tea shops more easily, Toa Payoh has long history of HDB housing, with very good infrastructure and facilities. As a result, it is a good choice for citizens.

Near downtown Singapore, there are moderate numbers of bubble tea shops, which is **cluster 0**. Investors may consider open bubble tea shops there, but the rental fee is high and there is still competition.

Most consumers in downtown area is tourists and people want to shop, so the investors may consider the impacts of volume of commuters.

PAs in **cluster 1** is a little far away from downtown area, and with low number of bubble tea shops. Those areas are mostly residential places for citizens. Investors may consider to open bubble tea shops near neighbourhoods to have more customers. The rental fee is relatively low, so the cost can be greatly reduced. **Cluster 1** seems to be the most suitable places to open bubble tea shops.

6. Conclusion

In this project, we have using python, along with other machine learning and visualization strategies to analyse the PAs in central Singapore on bubble tea shop opening problems. We gain data from Wikipedia, Geocoder and Foursquare to build the data frame.

We cluster PAs into 3 groups based on the frequency of the bubble tea shops and give suggestions to the investors. We have found that Toa Payoh has most bubble tea shops, which may appeal some bubble tea lovers. Downtown areas have moderate shops, while residential areas have low number of shops. We recommend investors to open shops in cluster 1, which is residential areas, to avoid high rental fee and competition between different bubble tea shops.

7. Further improvement

While choosing location is one of the most vital part in opening a bubble tea shop, there are many other impacts we can further investigate in the future using data analysis.

For example, we can investigate the brands of bubble tea shops: which one is the most popular, which has the highest number of visits, and which has the highest rating.