IBM COURSERA

The Food Project of Singapore by Troy Lee

The Battleground of The Neighbourhood

Table of Contents

Introduction

Singapore is famous for its cleanliness, garden city and efficient government. It welcomed 18.5 million tourists into the country in 2018. However, what is less known is that it is also a food paradise, where all kinds of food can be found, due to its cosmopolitan city nature. What is most important, at least for me is that, I can get all of these delicious food for cheap.

There is so much food to eat in Singapore, and for tourists without enough time in Singapore, they might miss out!

The objective of this project is to help tourists find the most authentic yet diverse food in Singapore if they only have 3 days in Singapore.

That would mean on average 9 meals. Certain tourists have more disposable income to spend, and other tourists, like me, will try to eat the most local and cheap food when I travel. This recommendation will target both sets of tourists. The hawker centre locations will recommend more diverse, delicious and cheap food, but sometimes reviews of its awesomeness are not online. However, most of the reviews of restaurants (different from hawker centres) can be found online. I refuse to believe that the more expensive food will mean better quality. Quality of local food means authenticity for me, it is high quality food if I find many locals queuing up for the food.

Data

The Foursquare API will be used to pull the following location data on restaurants in Singapore:

- *Venue Name*
- *Venue ID*
- *Venue Location*
- *Venue Category*
- *Count of Likes*

To acquire the data mentioned above, the following is done:

- Get geolocator latitude and longitude coordinates for Singapore
- Use Foursquare API to get a list of all venues in Singapore
- Get venue name, venue ID, location, category, and likes
- Obtain hawker centres locations from Singapore's open data: https://data.gov.sg/dataset/list-of-government-markets-hawker-centres

The assumption is that the number of likes is proportionate to the quality of food.

However, quality of food can be based on so many factors such as but not limited to atmosphere, price, convenience, fame, view, cleanliness, variety of food, service quality, music / ambience, aroma, weather of that day, taste preferences, audience who voted (tourists/locals) and many more. It is difficult to ascertain that if many people love the food, you will love it too. There are many variables to be considered, however, due to the point of this project. Let us assume that the number of likes of indeed proportionate to the quality of the food.

We are also assuming that this sample represents the whole of Singapore, which is untrue. Many amazing restaurants might not be voted by its lovers. Some traditional restaurants might have websites yet (I know, I know, you are crying out 'We are in the digital age, are these people so cavemen! That they don't even have a website.) You have to understand that some of these amazing restaurants in Singapore might still be run by people who are

above 60 or 70 years old, yes they might have an iPhone and play Candy Crush at home, but they might not see the need for a website. Sometimes, for simple reasons, they are so famous locally that they cannot even handle the normal onslaught of customers, they would like to sustain the existing number of customers and not get multitudes more of customers. Hence this sample does not represent the whole of Singapore.

Methodology

The gathered data will be used to create a k-means clustering algorithm that groups restaurants into 4 clusters so that people looking to eat in Singapore can easily see which restaurants are the best to eat at, what cuisine is available and the location. Some pictures of the food were introduced and I have left comments if I have eaten there previously.

The geographic data was found for Singapore using the FourSquare API. After using the coordinates to obtain a group of organisations, they were filtered. Any organisations that did not involve food were removed.

The location of the remaining restaurants were obtained.

Next, the environment was set up to obtain the number of likes from these restaurants from the API based on the venue ID.

Further grouping of the restaurants based on the amount of likes they had was carried out.

Dummy variables were created to allow for clustering.

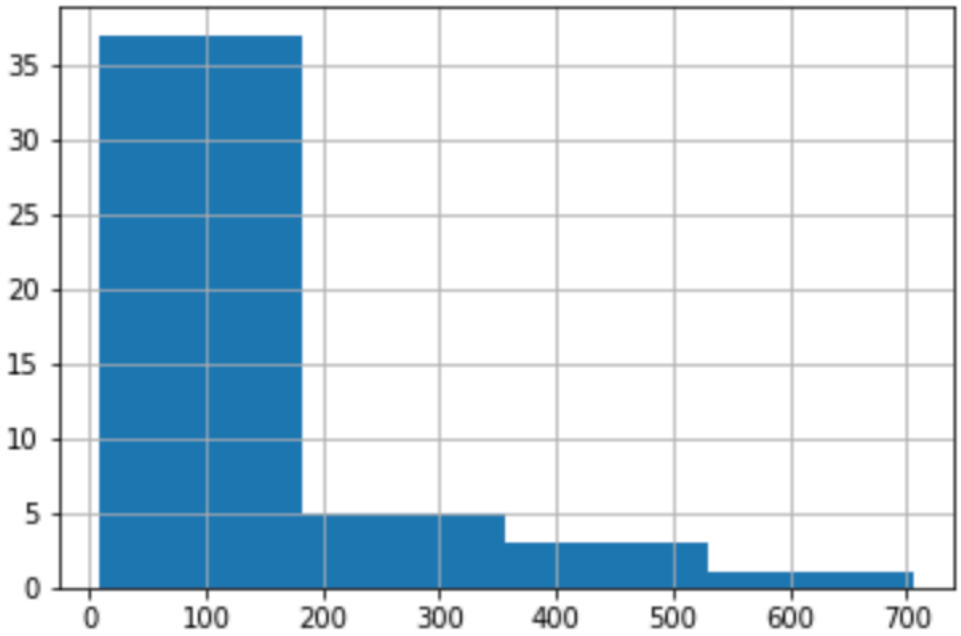
K-Means algorithm was run to obtain the clusters. The clusters were then added back to the original dataframe.

A map was created with markers to visualize the locations of these restaurants.

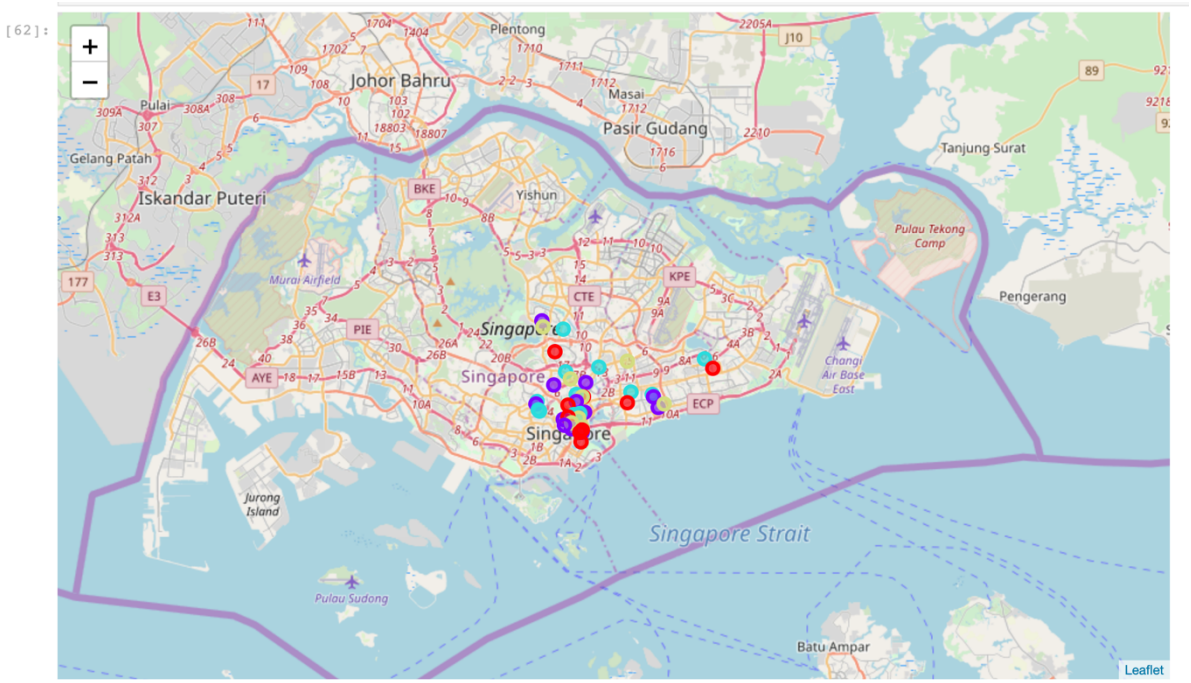
Lastly, hawkers centres were introduced.

Results

The distribution of the number of likes of the restaurants are shown below.



The following depicts the locations of the restaurants on the maps. They are all located at the South of the island.



Cluster 1 to 4, depict the restaurants with decreasing amount of like, are shown below.

Cluster 4 is not shown.

Cluster 1

```
[63]: sg_venues.loc[sg_venues['label']==0]
```

[63]:

	name	id	categories	lat	lng	total likes	total likes_cat	categories_new	label
7	Old Airport Road Food Centre	4b0bcea6f964a520c93323e3	Food Court	1.308171	103.885811	705	great	Others	0
9	Chye Seng Huat Hardware Coffee Bar	50227b21e4b02e5c64efc45d	Coffee Shop	1.311675	103.860374	527	great	Others	0
11	Creamier Ice Cream And Coffee	4e9fb0b46c258f17cb31f906	Ice Cream Shop	1.337411	103.844086	227	great	Others	0
12	Druggists	54ccdb92498eeb9f4153347e	Bar	1.311227	103.859830	123	great	Bars	0
18	ATLAS Bar	585a1ef68ee5606a98d90f21	Cocktail Bar	1.299965	103.857768	193	great	Bars	0
23	Cicheti	51fb8ad5498ec467c6e817e4	Italian Restaurant	1.301707	103.860062	112	great	European Cuisine	0
29	Komala Vilas Restaurant	4b058813f964a5204cb022e3	Vegetarian / Vegan Restaurant	1.306523	103.851848	141	great	Others	0
33	Tom's Palette	4bd811e909ecb7133ffc477c	Ice Cream Shop	1.296079	103.856757	147	great	Others	0
52	edge | food theatre	5052db48e4b0359ed4970086	Buffet	1.291870	103.859241	103	great	Others	0
58	Artichoke Café + Bar	4c6f6103df6b8cfa1342b54d	Middle Eastern Restaurant	1.299678	103.851978	210	great	Middle East Cuisine	0

Cluster 2

```
[64]: sg_venues.loc[sg_venues['label']==1]
```

[64]:

	name	id	categories	lat	lng	total likes	total likes_cat	categories_new	label
2	Boon Tong Kee 文东记	4b9b68a5f964a5208a0536e3	Chinese Restaurant	1.319819	103.862028	70	abv avg	Asian Cuisine	1
17	Loving Hut	52d0b63811d2fd7c68dde58d	Vegetarian / Vegan Restaurant	1.311118	103.901190	69	abv avg	Others	1
20	Birds of Paradise Gelato Boutique	5794bbb9498eaba989965ceb	Ice Cream Shop	1.305051	103.903787	67	abv avg	Others	1
21	Murugan Idli Shop	4b2c6299f964a52088c624e3	Indian Restaurant	1.308842	103.856380	74	abv avg	Asian Cuisine	1
24	Ah Chew Desserts 阿秋甜品 (Ah Chew Desserts)	53492538498e0f29cc1f4da9	Dessert Shop	1.318411	103.843714	49	abv avg	Others	1
27	Sin Ming Roti Prata (Faisal & Aziz Curry Musli...	4b584116f964a520c14f28e3	Indian Restaurant	1.355269	103.836719	75	abv avg	Asian Cuisine	1
28	Sari Ratu Restaurant	4fdaf746e4b0bcab80d803b7	Indonesian Restaurant	1.302145	103.861258	47	abv avg	None	1
47	The Auld Alliance	4db1a2035da32cf2df60580b	Whisky Bar	1.298567	103.848903	58	abv avg	Bars	1
50	Colony	55f9641d498eaceb29f0e5bb	Buffet	1.290697	103.859914	63	abv avg	Others	1
74	JAAN	4b058813f964a5203ab022e3	French Restaurant	1.293161	103.853583	75	abv avg	European Cuisine	1

Cluster 3

```
[65]: sg_venues.loc[sg_venues['label']==2]
```

[65]:

	name	id	categories	lat	lng	total likes	total likes_cat	categories_new	label
1	Ser Seng Herbs (Turtle) Restaurant 生成山瑞補品	4bc2ebfc4cdfc9b645a49621	Chinese Restaurant	1.328688	103.869835	30	below avg	Asian Cuisine	2
4	Tandoori Corner	4b3af9cbf964a520197025e3	Indian Restaurant	1.325778	103.850011	36	below avg	Asian Cuisine	2
5	Sik Bao Sin (Desmond's Creations)	525a573711d2af3d09ec3024	Cantonese Restaurant	1.314241	103.887825	29	below avg	Asian Cuisine	2
8	Dian Xiao Er 店小二	54017053498ed9078533b8cd	Chinese Restaurant	1.350426	103.848988	13	poor	Asian Cuisine	2
14	MTR 1924	51a2af4d498ec025802f527d	Indian Restaurant	1.312973	103.856371	91	great	Asian Cuisine	2
15	Long Phung Vietnamese Restaurant	4b3dd5fbf964a520039725e3	Vietnamese Restaurant	1.312703	103.900227	195	great	Asian Cuisine	2
44	Singapore Zam Zam Restaurant	4b86050ef964a520b17d31e3	Indian Restaurant	1.302030	103.858548	526	great	Asian Cuisine	2
46	Ramen Keisuke Tonkotsu King Four Seasons	5187c434498e1a1f7d3537b2	Ramen Restaurant	1.301017	103.855381	176	great	Asian Cuisine	2
55	Seng Kee Black Chicken Herbal Soup Bedok North...	4c613f0e13791b8d956051af	Asian Restaurant	1.333087	103.930373	11	poor	Asian Cuisine	2
56	Tatsuya Japanese Restaurant	4b6ee15ff964a520c1ce2ce3	Japanese Restaurant	1.308391	103.833997	96	great	Asian Cuisine	2
87	Din Tai Fung 鼎泰豐 (Din Tai Fung)	4b7152c1f964a5201e412de3	Dumpling Restaurant	1.303350	103.835031	391	great	Asian Cuisine	2

The following shows an ending recommendation for tourists to visit local hawker centres and not just restaurants if they want to live in a more local way.

[79]:
```
df = df.sort_values(['no_of_stalls'], ascending = False)
df.head(20)
```

[79]:

	name_of_centre	location_of_centre	type_of_centre	owner	no_of_stalls	no_of_cooked_food_stalls	no_of_mkt_produce_stalls
100	Chinatown Market	Blk 335, Smith Street, S(050335)	MHC	HDB	703	226	477
106	Tekka Market	Blk 665, Buffalo Road, S(210665)	MHC	HDB	403	119	284
10	Geylang Serai Market	Geylang Serai, S(402001)	MHC	Government	365	63	302
23	Tiong Bahru Market	30, Seng Poh Road, S(168898)	MHC	Government	342	83	259
26	Albert Centre Market & Food Centre	Blk 270, Queen Street, S(180270)	MHC	HDB	234	86	148
40	Blk 20 Ghim Moh Road	Blk 20, Ghim Moh Road, S(270020)	MHC	HDB	230	72	158
51	Blk 347 Jurong East Ave 1	Blk 347, Jurong East Ave 1, S(600347)	MHC	HDB	225	56	169

Discussion

K-Means clustering has grouped the restaurants into likes as well as the areas. As we can see, it is inefficient in doing so. The picture below depicts three restaurants in cluster 3, however, the number of likes they have should have put them into cluster 1. Hence we cannot trust everything that the machine has segmented for us. No matter how much the metrics are changed, there will always be some errors. And this presents the need for human interpretation of the data before releasing to the audience.

If anyone is coming to Singapore, I would recommend them trying: Old Airport Road Food Centre, Zam Zam, Tong Le Private Dining (which is not on the list), Wing Seong's Fatty(which is not on the list), go to Chomp Chomp Serangoon Gardens at night (to experience the spot for supper for the local youth, again this is not depicted in the data).

Conclusion

Singapore has a lot of food to welcome tourists. However, they need to know what they are looking for. I personally would recommend tourists to try more local food than visiting the restaurants which serve cuisine that they are familiar with. However, they cannot trust all the reviews or recommendations online as well, as we have seen through this report, that some reviews might be inefficient in categorising quality. In addition, the small sample used

cannot be representative of all Singapore food. Even if a large sample is chosen, we need to understand that certain restaurants might not even be within the list of samples to be chosen, simply because they lack an online presence.

Therefore, the best experience is always to regard reviews with a pinch of salt, and ask locals where they would prefer to eat in the area, stating to them what you might be looking for. I believe that remains the best way to understand a country, not just Singapore. Talk to a local, and try authentic local food.