Food Trail in Singapore!

An IBM Data Science Capstone Project by Caden Lee

Introduction

Singapore is an island-state with 721.5 sq km landscape and a population of 5.6 million people (as of 2018). It is well-known for its multicultural and multiracial society, on top of being an Asian Financial Powerhouse and the City of the Future. Many tourists are drawn to Singapore to experience the multiculturalism. But more importantly, many come for the great food in Singapore. Singapore has a wide variety of food that originates from different parts of Asia and the world. Food is a common theme among the people and one of the most effective gel that holds the society cohesive.

This report shall explore the towns in Singapore and the food establishments. This finding can be interesting to tourists who wish to experience Singapore and its local culture through food. Also, this report shall address the ongoing argument that there is difference between East and West Singapore when it comes to food. Ultimately, the author wishes to provide readers an insight to the food in Singapore through data.

Data

The data used in this analysis are

1. Location data of the different MRT (Mass Rapid Transport) stations as references to the towns in Singapore;
2. Foursquare data of the different food establishments found in each town.

The MRT station data is enough as the MRT system (the local subway system) is expansive and covers most of the towns in Singapore. Most housings are built around the MRT and hence there are many food establishments around it. The data is taken from github/xkjyeah/singapore-postal-codes.

The Food Establishment data is extracted from Foursquare via its API.

Methods

1. Data Visualization

Python Folium package is used to visualize the data on the Singapore map.

2. Data Analysis

In this analysis, the data analysed are mainly categorical data. Hence, hot encoding is used to create dummy variables. As the objective of the exercise is to spot differences of the different neighbourhood, clustering is used as the machine learning algorithm.

Results

A map of Singapore marked with the different clusters can be found in Figure 1. Using the clustering algorithm, the towns were split into three distinct groups.

Cluster 0 consists of mainly Food Courts, Coffee Shops and Fast Food Restaurants, as seen in Figure 2. Coffee shops and Food Courts are centralized food establishments. One may find many kinds of food in food courts, such as Chinese, Malay, Indian, Western and Local Delights. While the food sold in food courts is relatively inexpensive, the food tastes delicious.

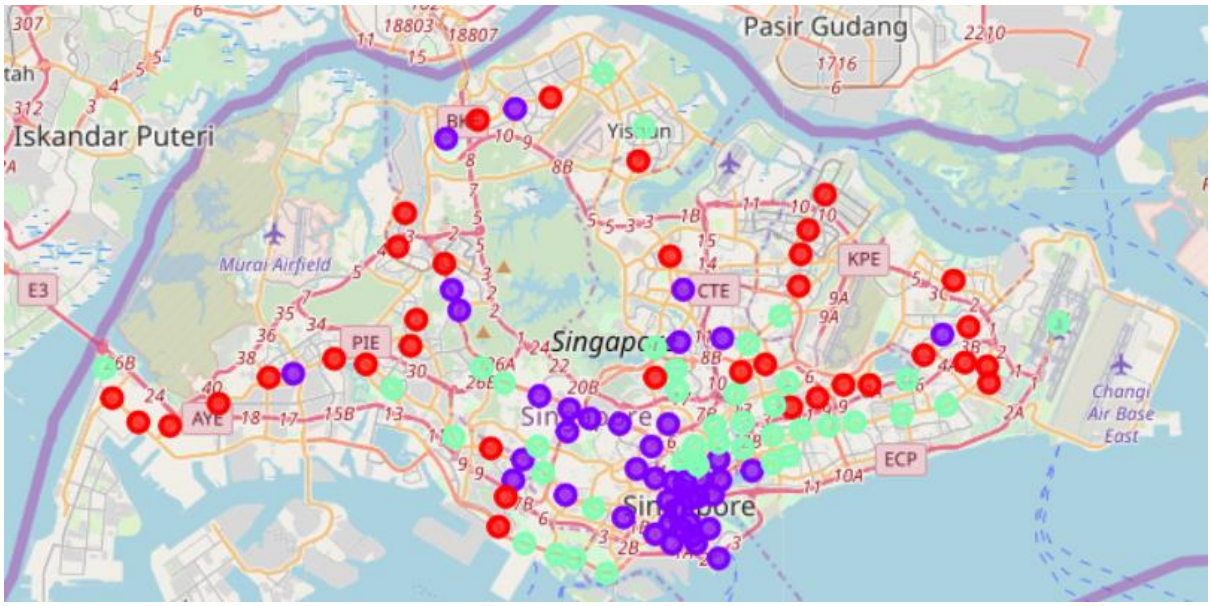

Figure 1. Clusters on the Singapore map. Red = 0, Purple = 1, Green = 2.

	STATION NAME	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
1	Bukit Batok	0	Food Court	Coffee Shop	Chinese Restaurant	Malay Restaurant	Fast Food Restaurant
2	Bukit Gombak	0	Food Court	Coffee Shop	Chinese Restaurant	Fast Food Restaurant	Indian Restaurant
3	Choa Chu Kang	0	Fast Food Restaurant	Coffee Shop	Asian Restaurant	Thai Restaurant	Food Court
4	Yew Tee	0	Fast Food Restaurant	Food Court	Café	Japanese Restaurant	Restaurant
6	Marsiling	0	Food Court	Coffee Shop	Café	Fast Food Restaurant	Asian Restaurant

Figure 2. Cluster 0

Cluster 1 towns have restaurants that are generally more exotic, such as Japanese, European, Korean and even Cafés. Looking at the type of food establishments, the food sold in towns are likely to be more expensive (Figure 3).

Except for Kranji, the towns marked as cluster 1 are relative central. Tampines and Woodlands are exceptions also, but they are commercial hubs where many office buildings are located.

17	Novena	1	Coffee Shop	Vegetarian / Vegan Restaurant	Italian Restaurant	Café	Chinese Restaurant
18	Newton	1	Japanese Restaurant	Chinese Restaurant	Spanish Restaurant	Seafood Restaurant	Thai Restaurant
19	Orchard	1	Sushi Restaurant	Chinese Restaurant	Asian Restaurant	Indonesian Restaurant	Restaurant
20	Somerset	1	Japanese Restaurant	Ramen Restaurant	Chinese Restaurant	Korean Restaurant	Dumpling Restaurant
21	Dhoby Ghaut	1	Café	Hotpot Restaurant	Korean Restaurant	Modern European Restaurant	Indonesian Restaurant

Figure 3. Cluster 1

Most cluster 1 towns are in the downtown, where the Central Business District and the Tourist Hotspots are. Hence it is expected that food variety is greater and costs more expensive in these areas.

	STATION NAME	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Jurong East	2	Chinese Restaurant	Japanese Restaurant	Italian Restaurant	Halal Restaurant	Korean Restaurant
9	Sembawang	2	Chinese Restaurant	Coffee Shop	Asian Restaurant	Fast Food Restaurant	Café
10	Yishun	2	Indian Restaurant	Chinese Restaurant	Thai Restaurant	Coffee Shop	Vegetarian / Vegan Restaurant
15	Braddell	2	Chinese Restaurant	Asian Restaurant	Coffee Shop	Japanese Restaurant	Seafood Restaurant
16	Toa Payoh	2	Chinese Restaurant	Coffee Shop	Indian Restaurant	Asian Restaurant	Hong Kong Restaurant

Figure 4. Cluster 2

Cluster 2 and Cluster 1 towns are alike. However, the main difference is that the food establishments are more local, i.e. Chinese, Indian, Malay and Asian restaurants are more commonly found (Figure 4). In other words, the restaurants in cluster 2 towns are more local compared to Cluster 1.

Discussion

From the map in Figure 1, one may observe that clusters of the same type may be spread into different area and may form separate sub groups. To differentiate the different areas (North, East, West and Central), instead of looking at individual cluster, one may look at the cluster density in each area.

East has the highest cluster 2 density compared to other areas while Central has the highest cluster 1 density. West and North are more similar as both are more populated by Cluster 0. With this, one may conclude that East and West Singapore are indeed very different, especially when it comes to the type of food establishments in the area, and the inferred price tag of the food.

To improve on this report, one may look further into the average true price of the food sold in each area. Also, to provide a more holistic view of the differences among the areas, one may investigate other factors such as recreations, education and commercial.

Conclusion

This report has briefly explored the differences among the areas in Singapore, more particularly between East and West Singapore. The finding of this report is that East Singapore generally has more Asian restaurants compared to other areas. West Singapore has more food courts than restaurants compared to the rest.