

The best area to open a food or drink business in Singapore development areas.

Introduction

In this report we will attempt to find the most popular venue locations in the Singapore development areas of Bedok, Changi, Changi Bay, Pasir Ris, and Paya Lebar to open a food or drink business such as a restaurant or coffee shop.



This information that we shall be investigating will help make decisions for areas of business development but will not specifically suggest what type of business would be best to start in a certain area.

Data Sources

The 2 main sources of data will be foursquare api for gathering information of the top venues in each area, and Wikipedia for gathering population and mapping coordinates of each development area. The data that will be used will be mainly coordinates for each area, number of different types of places associated with each area and population for referencing against the findings.

Methodology

In this investigation we used a variety of data analysis and exploration techniques in order to present the findings.

To begin we used pandas read_html for gathering the data set from wikipedia. This was a useful tool as it allowed for a datatable on the wikipedia website to be pulled into a dataframe on the python notebook.

We then used the pandas library to investigate and clean up the data set, removing the columns that we would not be using and disregarding any missing data for the investigation.

BEFORE CLEAN UP

	Neighborhood	Malay	Chinese	Pinyin	Tamil	Region	Area	Population	Density
0	Bedok	*	勿洛	Wù luò	பிடோக்	East	21.69	279380	13000
1	Bishan	NaN	碧山	Bì shān	பீஷான்	Central	7.62	88010	12000
2	Boon Lay	NaN	文礼	Wén lǐ	பூன் லே	West	8.23	30	3.6
3	Bukit Batok	*	武吉巴督	Wǔjī bā dū	புக்கிட் பாத் தோக்	West	11.13	153740	14000
4	Bukit Merah	*	红山	Hóng shān	புக்கிட் மேரா	Central	14.34	151980	11000

AFTER CLEAN UP

	Neighborhood	Region	Area	Population	Density	Latitude	Longitude
0	Bedok	East	21.69	279380	13000	1.323608	103.927340
8	Changi	East	40.61	1830	80.62	1.345010	103.983209
9	Changi Bay	East	1.70	*	*	1.321971	104.029002
29	Pasir Ris	East	15.02	148020	9600	1.372094	103.947373
30	Paya Lebar	East	11.69	40	3.4	1.351609	103.899516

Now we have a usable dataset in order to use with the foursquare api. It contains all the necessary data for our investigation to be continued.

Using the foursquare api we searched for all the venues associated with the locations and listed them in a separate dataframe.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Bedok	1.323608	103.92734	Ya Kun Kaya Toast 亞坤	1.324095	103.929198	Coffee Shop
1	Bedok	1.323608	103.92734	FairPrice Finest	1.324140	103.929260	Supermarket
2	Bedok	1.323608	103.92734	Haidilao Hot Pot 海底捞火锅	1.324299	103.929104	Hotpot Restaurant
3	Bedok	1.323608	103.92734	Bedok Chwee Kueh 勿洛水糰	1.324903	103.930250	Chinese Restaurant
4	Bedok	1.323608	103.92734	Song Zhou Luo Bo Gao 松洲萝卜糕	1.324836	103.930520	Breakfast Spot

This data frame above is the example of the head of the new data frame that we created with the new location venues listed. From here we can now count the number of each unique venues within each location.

Neighborhood						
Bedok	61	61	61	61	61	61
Changi	12	12	12	12	12	12
Changi Bay	5	5	5	5	5	5
Pasir Ris	25	25	25	25	25	25
Paya Lebar	2	2	2	2	2	2
Tampines	8	8	8	8	8	8

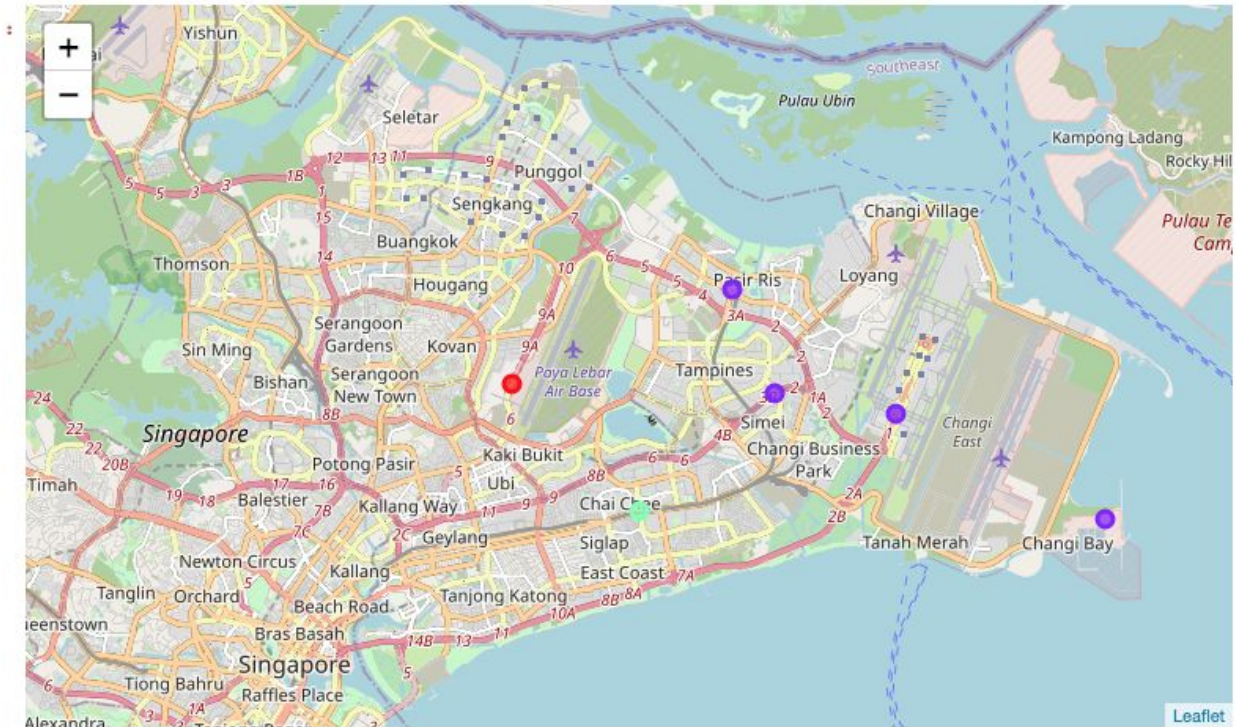
This dataframe above tells us how many different types of venues there are in each development location. Using this data frame we can start to see which are the most popular categories in each location.

After this stage we started to group the venues in terms of population so that we can describe which are the most frequently occurring venue in each given location. To achieve this we went through stages of grouping and hotcoding.

	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
2	Changi Bay	Boat or Ferry	Military Base	Wings Joint	Chinese Restaurant	Frozen Yogurt Shop	Fried Chicken Joint	French Restaurant	Food Court	Flower Shop
1	Changi	Bus Station	Airport Terminal	Road	Coffee Shop	Café	Men's Store	Movie Theater	Airport Service	Tunnel
0	Bedok	Coffee Shop	Chinese Restaurant	Sushi Restaurant	Supermarket	Café	Food Court	Bakery	Malay Restaurant	Noodle House
3	Pasir Ris	Food Court	Fast Food Restaurant	Supermarket	Coffee Shop	Sandwich Place	Outdoor Supply Store	Flower Shop	Diner	Chinese Restaurant
4	Paya Lebar	Military Base	Bus Station	Wings Joint	Chinese Restaurant	Frozen Yogurt Shop	Fried Chicken Joint	French Restaurant	Food Court	Flower Shop
5	Tampines	Sandwich Place	Grocery Store	South Indian Restaurant	Housing Development	Indian Restaurant	Café	Fast Food Restaurant	Plaza	Dessert Shop

The above data frame is an example of the data after grouping into most common venues in each development location. To use this data further with a folium maps we shall create cluster of the most populated areas of business.

After the clustering in complete we can plot the findings onto a map using folium maps.



The map above shows our clustered areas and using color coordination we can specify the levels of high populated businesses in each area.

Results

Our findings have shown us that the most business populated area is the bedok area (green colour on map) This area is located in a high populated area which shows that there is opportunity for a wide variety of businesses to grow and succeed in that area.

	Neighborhood	Region	Area	Population	Density	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
0	Bedok	East	21.69	279380	13000	1.323608	103.927340	2	Coffee Shop	Chinese Restaurant	Sushi Restaurant
8	Changi	East	40.61	1830	80.62	1.345010	103.983209	1	Bus Station	Airport Terminal	Road
9	Changi Bay	East	1.70	*	*	1.321971	104.029002	1	Boat or Ferry	Military Base	Wings Joint
29	Pasir Ris	East	15.02	148020	9600	1.372094	103.947373	1	Food Court	Fast Food Restaurant	Supermarket
30	Paya Lebar	East	11.69	40	3.4	1.351609	103.899516	0	Military Base	Bus Station	Wings Joint

From our data frame above we can see that Bedok has a high civilian population which agrees with our business population finding that it is the most highly business populated location of the set. We can also see that the 1st most popular venue in the Bedok area is coffee shops, and the 2nd most popular venue is chinese restaurants, and the 3rd most popular venue is sushi restaurants.

This data also suggests that the area of Pasir Ris which has the 2nd largest population from the development areas is also a good area of development for restaurants and supermarkets. They show that the 1st most popular venue in this area is food courts, the 2nd most popular is fast food restaurants and the 3rd is supermarkets.

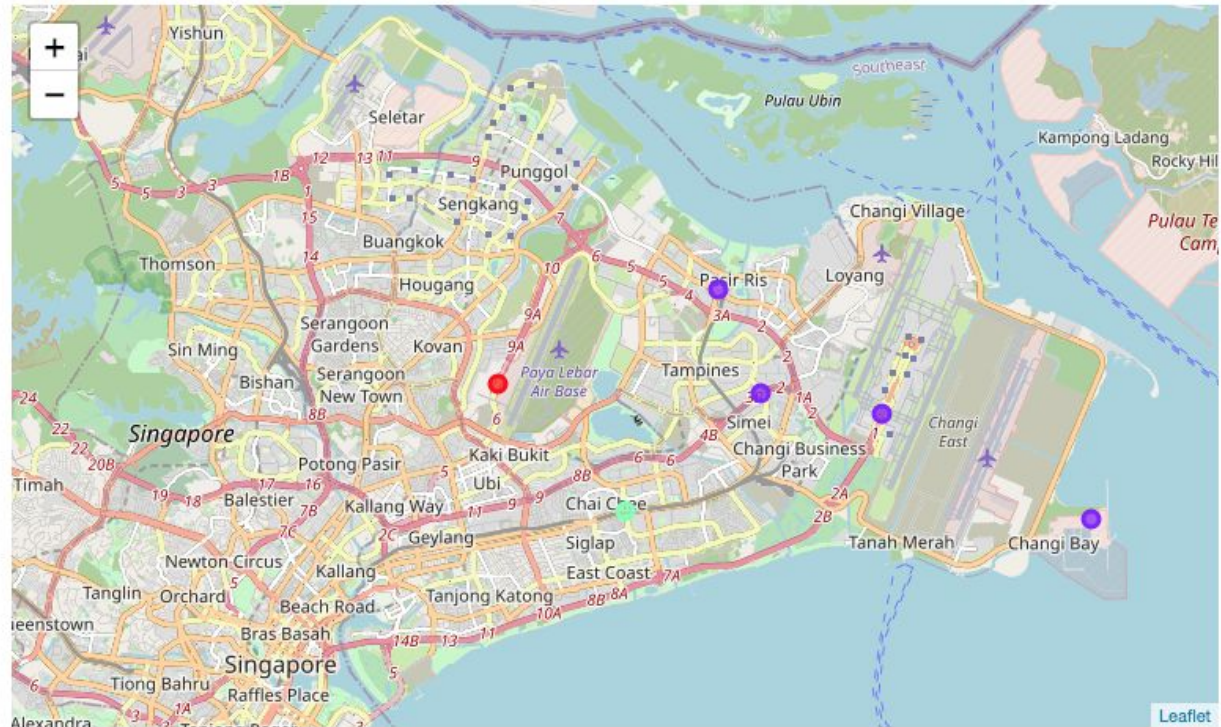
The areas of Changi, Changi Bay, and Paya Lebar all have a small population in common and do not seem to be residential areas but more industrial, transportation and military areas. The most popular food venus in these areas seem to be wing joints.

By looking at the cluster map below we can see the location of the development areas and make more sense of the data table above.

Red = Paya Lebar

Green = Bedok

Purple = Changi, Changi Bay, Pasir Ris



Discussion

Looking at the investigation and the results we can see a big difference in the development areas in terms of population which would suggest areas that are more residential than other which seem to be more industrial.

Bedok has the highest area of population and all of its top 3 venues are coffee shops or restaurants. This is a good indication that there is enough population to sustain these types of businesses and if a gap in the restaurant market can be found for this area, may present a good opportunity for a sustainable business. Although due to the already high population of restaurants and coffee shops in the area this could also mean there is a lot of competition in this area.

My best recommendation from this investigation is to start a business in the area of Pasir Ris as it has a good large population but also is not over-saturated by coffee shops and restaurant businesses. In my opinion I believe that it would be the best option to open a coffee shop in the area of Pasir Ris. By looking at Bedok as an example it shows that coffee shops are very

popular in the area and this seems to be missing from the Pasir Ris area which has similar statistics.

Conclusion

To conclude this report we would say that the best development area of Singapore would be the area of Pasir Ris. we came to this conclusion by analysing statistical data gathered from the foursquare api which told us the most popular venue types in the different areas. We used the high business successful and populated area of Bedok as an example of indication for future business opportunities.

The remaining areas that have not been discussed much in detail such as Changi, Changi Bay and Paya Lebar would not make a good option to open a food or drink business due to their low residential population.