Title: Singapore next CBD?

1. Introduction

Singapore is a city state with a relatively small land mass. Consequently, urban planning becomes a difficult issue. Any lack of planning thereof may easily snowball into a serious problem where mitigation may be extremely costly.

The Central Business District (CBD) is the main business and commercial centre of a town or city. In Singapore, the CBD is located along the south eastern part of the central area of Singapore within the downtown core district. Like many CBDs around the world, geographical reasons played an important role in determining the success of Singapore's CBD-It is relatively near the port area due to historical reasons, where convenient transport allowed business to flourish rapidly, which in turn has resulted in it being the main focus of any future developments surrounding those areas, allowing it to become what it is today.

However, with the ever-increasing population and to ensure fair development between the different areas in Singapore, as part of the urban redevelopment strategy, Jurong East has been selected and to be redevelopment as a potential "second" Central Business District in the country.

https://www.ura.gov.sg/corporate/media-room/media-releases/pr16-45

While Singapore may be a small country, if the plan above is successful, many years down the road, there may be a need to establish another CBD in Singapore. However, even then, as planning may take years to come into fruition. It is useful to plan, and to look for other potential regions in which new CBD areas may potentially be set up, so that planners can hit the ground running. Government officials may also need be able to use this understanding to convince the local population when the need eventually arises, since the transparency of a data-driven

approach may be able to better explain it to the local population when the need comes. This project aims to provide a potential solution to the problem discussed here.

2. Data description

Due to the space constrain, public transport is extremely popular in Singapore. The Mass Rapid Transit trains is relied upon by most people to get from point A to point B. Thus, any businesses that surrounds such areas may be useful in helping to determine whether a CBD should exist in the area.

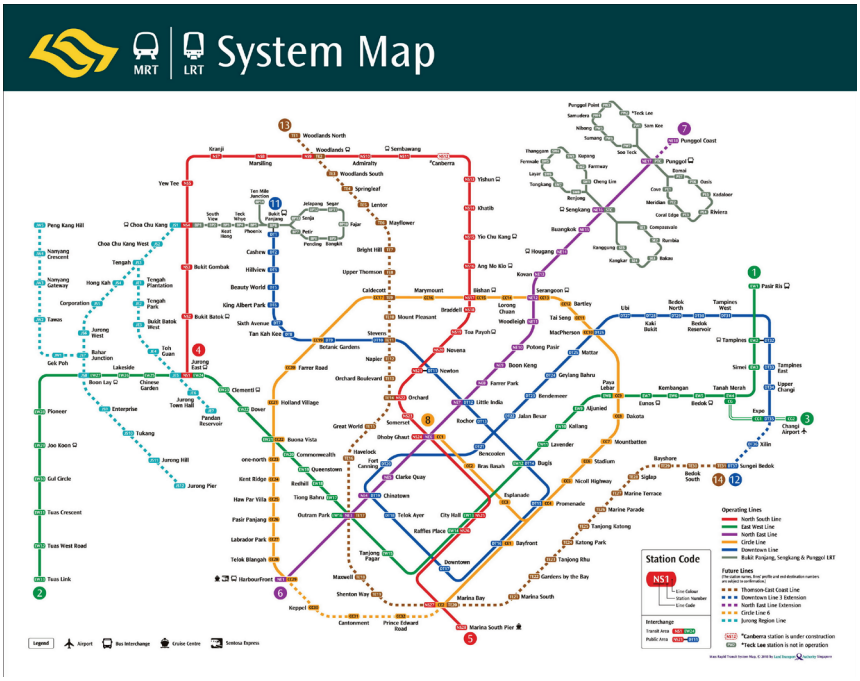

Figure 1 Singapore MRT System map

(https://www.transitlink.com.sg/images/eguide/mrt_sys_map.htm)

In order to obtain the necessary data, a list of MRT stations in Singapore is collected and compiled from Wikipedia, together with the corresponding GPS coordinates for each of the MRT stations.

[2]:

	Name	Lat	Lon	Abbrv
0	Jurong East	1.333415	103.742119	JUR
1	Bukit Batok	1.349073	103.749664	BBT
2	Bukit Gombak	1.358702	103.751787	BGB
3	Choa Chu Kang	1.385092	103.744322	CCK
4	Yew Tee	1.396986	103.747239	YWT

Figure 2 List of MRT stations with the corresponding names

The GPS coordinates of each of the stations are then fed into Foursquare's API to obtain the list of business, together with the corresponding business type that exists in the area. A search radius of 5km is set for each of this MRT stations, since it is highly unlikely for people to walk beyond this distance in a day to day scenario.

[7]:

	Neighbourhood	Neighbourhood Latitude	Neighbourhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Type
0	Jurong East	1.333415	103.742119	Jurong East MRT Interchange (NS1/EW24)	1.333503	103.742185	Train Station
1	Jurong East	1.333415	103.742119	Jem	1.333319	103.743117	Shopping Mall
2	Jurong East	1.333415	103.742119	Jurong East MRT Platform (NS/EW)	1.333311	103.742223	Platform
3	Jurong East	1.333415	103.742119	Monster Burger	1.333487	103.741796	Burger Joint
4	Jurong East	1.333415	103.742119	Westgate	1.334225	103.742824	Shopping Mall
5	Jurong East	1.333415	103.742119	Jurong East Temporary Bus Interchange	1.333259	103.741415	Bus Station
6	Jurong East	1.333415	103.742119	Kopitiam	1.333278	103.741619	Cafeteria
7	Jurong East	1.333415	103.742119	Kodak Express Photo-shop	1.333226	103.742151	Miscellaneous Shop
8	Jurong East	1.333415	103.742119	KOPI	1.333494	103.741682	Food Court
9	Jurong East	1.333415	103.742119	Jollibee	1.333470	103.741659	Fast Food Restaurant
10	Jurong East	1.333415	103.742119	Posb	1.333649	103.742143	ATM
11	Jurong East	1.333415	103.742119	BreadTalk	1.333495	103.743553	Bakery
12	Jurong East	1.333415	103.742119	Johan Paris	1.334083	103.742384	Bakery
13	Jurong East	1.333415	103.742119	Zekko	1.333206	103.742005	Shoe Store

Figure 3 Example of the data obtained from Foursquare

The mean for the number of each type of business near a specified MRT station is then obtained and categorised using a K-Means approach. The results are then plotted using folium and observed in order to see any trend.

3. Methodology

A K-Means Classification approach was selected as the main data since the aim is to find any observable trends in the distribution of the business, selected based on their proximity towards

the nearest MRT station. The categorisation is done based on a cluster size of 5. As the location of the existing CBD, and the upcoming CBD is known, we can use this as the basis of determining whether the K-Means classification approach can act as a suitable benchmark in determining whether the area is suitable to be established as a CBD.

There are two ways in which the data may be treated.

1. To see similarities between the data based on grouping, and to select the next CBD based on this classification

2. To provide suggestions on the next CBD based on the other categorisation, and to suggest possible improvements in the area.

The data set that is obtained using Foursquare's search function is incomplete, since many small businesses are labelled as "None". However, it is assumed that being labelled "none" means that the data may not be sufficiently defined and is dropped during the data processing stage. The final dataset consists of approximately 18000 data points. (see Notebook file for details)

98	Jurong East	1.333415	103.742119	Yolé	1.334004	103.742329	Frozen Yogurt Shop
99	Jurong East	1.333415	103.742119	On 78	1.334279	103.742506	None

Figure 4Example of a "None" dataset.

Within each neighbourhood, the 10 most popular types of business are then ranked. It is based on this data that a K-Means approach is then used.

```
----Admiralty----
             venue  freq
0       Food Court  0.04
1  Doctor's Office  0.03
2            Café  0.03
3    Train Station  0.03
4             Bank  0.03


----Aljunied----
                          venue  freq
0               College Classroom  0.08
1             Chinese Restaurant  0.08
2              Asian Restaurant  0.04
3                   Coffee Shop  0.03
4  Vegetarian / Vegan Restaurant  0.03


----Ang Mo Kio----
                venue  freq
0        Snack Place  0.05
1     Cosmetics Shop  0.04
2             Bakery  0.04
3  Mobile Phone Shop  0.03
4           Pharmacy  0.03
```

Figure 5 Example of frequency data

4. Results.

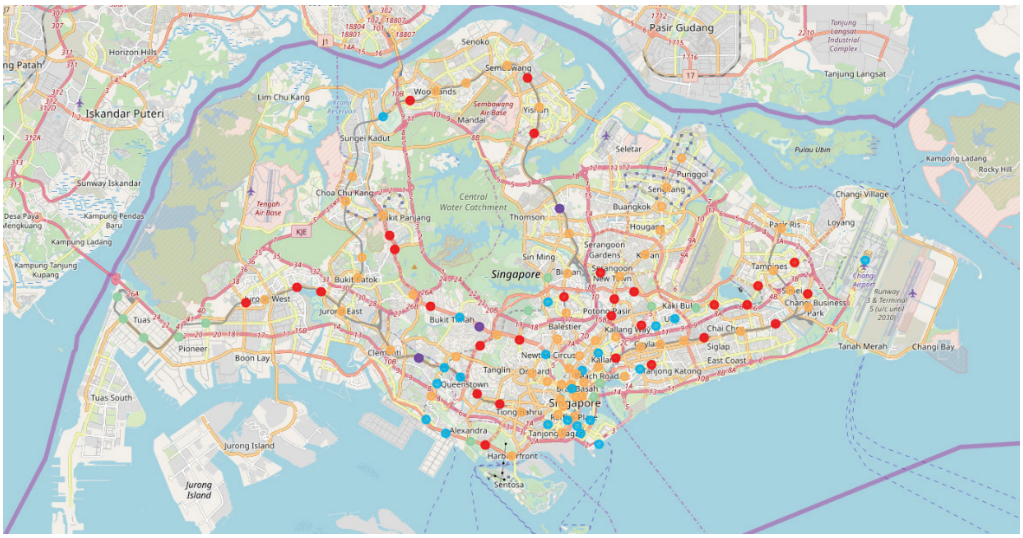

Figure 6 K means Result

Figure 1 shows the cluster results obtained using K-Means. The areas that are consistent with the downtown core (CBD) are all categorised under the same cluster (Cluster 4), which helps to prove the validity of the result

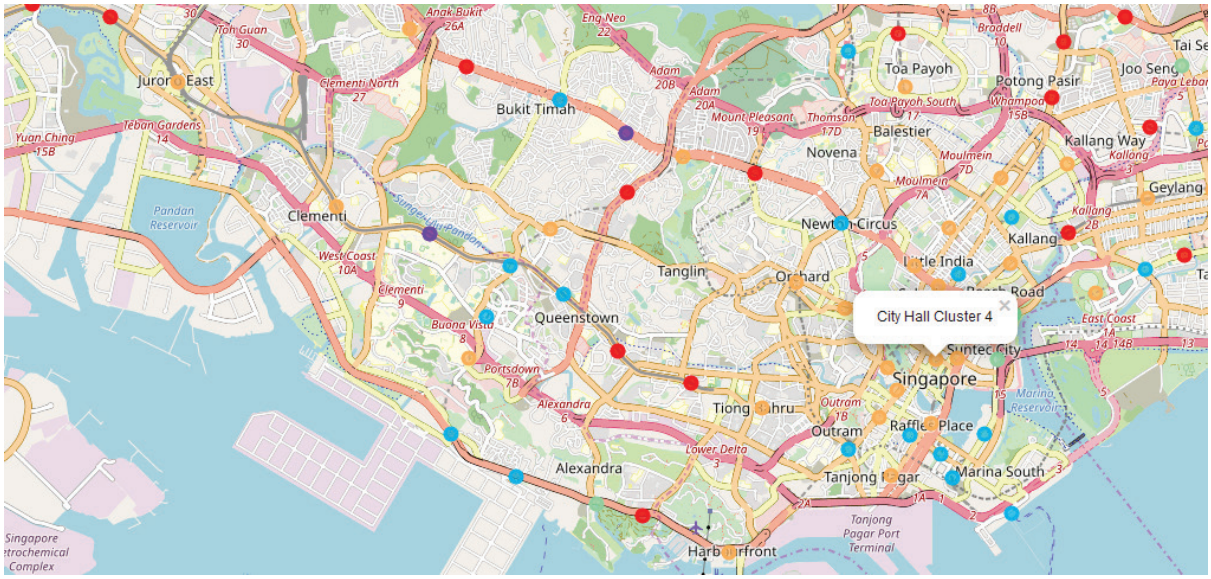

Figure 7 K Means for selected region of the map.

The most common venues for cluster 4, in which the CBD is also categorised as such, are listed in Figure 8 while the location of the next CBD (categorised as cluster in the data and labelled orange) has the following shop distribution as shown in figure 9.

[16]:

	Grouping	Abbrv	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
4	2	BFT	Event Space	Office	Jewelry Store
10	2	BDM	Office	Building	Café
16	2	BBS	College Classroom	Office	Café
23	2	BNV	Office	Miscellaneous Shop	Café
24	2	CDT	Flower Shop	Building	Office
27	2	CGA	Airport Terminal	Office	Gift Shop
34	2	COM	Building	Office	Residential Building (Apartment / Condo)
38	2	DTN	Office	Coffee Shop	Café
48	2	HPV	Office	Automotive Shop	Residential Building (Apartment / Condo)
52	2	JLB	Office	Hardware Store	Furniture / Home Store

Figure 8 Top 3 venues for area located within CBD

[18]:

	Grouping	Abbrv	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
0	4	ADM	Food Court	Parking	Bakery
1	4	ALJ	College Classroom	Chinese Restaurant	Asian Restaurant
2	4	AMK	Snack Place	Cosmetics Shop	Bakery
5	4	BTW	Chinese Restaurant	Café	Korean Restaurant
6	4	BDK	Noodle House	Dessert Shop	Cosmetics Shop
9	4	BCL	Café	Hotel	Bar
11	4	BSH	Food Court	Cosmetics Shop	Bakery
12	4	BNK	Chinese Restaurant	Noodle House	Coffee Shop
13	4	BNL	Bank	Bakery	Japanese Restaurant
14	4	BTN	Women's Store	Furniture / Home Store	Church
17	4	BGK	Dim Sum Restaurant	Chinese Restaurant	Coffee Shop

Figure 9 Supporting areas (labelled as orange)

5. Discussion

The data is shown to be consistent with the existing knowledge since the CBD area is grouped under the same category. However, based on this data, it contradicts the idea of setting up CBD in the next CBD location since it is not categorised under the same category.

However, the categorisation of the next CBD is consistent with the support services that are required for a CBD to thrive (e.g. restaurants etc), thus, consistent with the existing decision by the existing government, a CBD at that location can be concluded as ideal.

Looking at the data for the other groupings, from observation, it can be concluded that the other regions mainly consist of business that provide residential-type support, which may not be appropriate for a potential CBD location.

[16]:

	Grouping	Abbrv	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	Co
3	1	BLY	College Classroom	Residential Building (Apartment / Condo)	Bus Station	
7	1	BDN	Residential Building (Apartment / Condo)	Housing Development	Assisted Living	
8	1	BDR	Residential Building (Apartment / Condo)	Gym	College Lab	
15	1	BDL	Residential Building (Apartment / Condo)	Salon / Barbershop	Housing Development	
25	1	TBA	Residential Building (Apartment / Condo)	Asian Restaurant	Coffee Shop	Chin
26	1	CSW	Military Base	Residential Building (Apartment / Condo)	Bus Station	
29	1	CNG	Residential Building (Apartment / Condo)	Salon / Barbershop	Chinese Restaurant	

Figure 10 Example of the data from the other regions based on the grouping.

6. Conclusion

Based on the K-Means results, it can be concluded that the CBD area currently selected is consistent with the current understanding on how CBD's are formed. The same methodology can be used to select for potential CBD's in the future since there are also other areas that are categorised the same according to the KMM results.