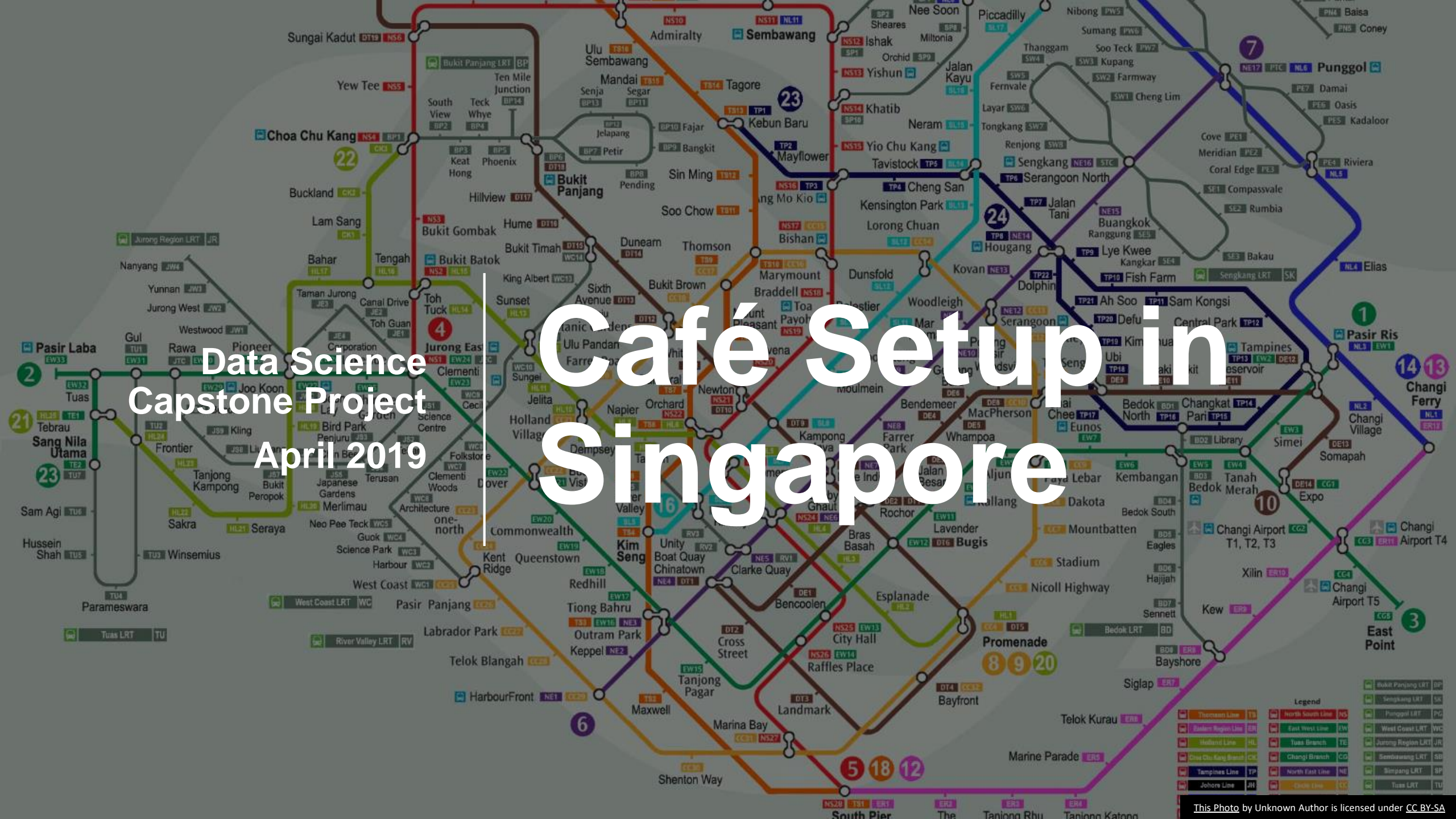


Data Science
Capstone Project
April 2019

Café Setup in Singapore



Problem Description and Background

- Singapore is known as a food haven, with many new eateries, restaurants and cafes springing up each year. However, the failure rate is also high. Some commonly cited statistics include:
- Of the 369 cafes, coffee houses and snack bars set up in 2011, over half had shut down by 2014
- Of the 391 cafes, coffee houses and snack bars registered in 2013, approximately a quarter had called it quits by the end of 2014
- A variety of reasons have been cited for these high failure rates, such as lack of experience, poor planning and a tight labour market. Location is another key factor that can affect business success.



Intended audience and problem refinement

Using FourSquare data, we hope to provide a budding entrepreneur who is hoping to start a cafe in Singapore with recommendations on where to set up their cafe. To ensure that:

- There will be adequate foot traffic (and in line with the urban decentralisation plan, e.g. <https://www.population.sg/articles/singapores-urban-planning-in-five-points>),
- Cafe is easily accessible

the focus will be on identifying which train stations will serve as the most suitable locations for setting up a cafe. Of particular interest is one of the newest train lines - the North-East (Purple) train line. This train line has 16 stations.

It is assumed that an ideal location for a cafe will be one that:

- Does not already have lots of cafes
- Has other eateries



Overview of datasets



1. Train station dataset

Created by user hxchua and uploaded on the dataworld website (<https://data.world/hxchua/train-stations-in-singapore>). The list of train stations was compiled in June 2017. It contains 9 variables and the ones of primary interest for this project are:

- Station Name
- Latitude
- Longitude

2. FourSquare dataset

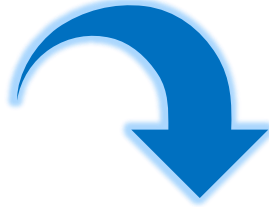
The extraction was based on the coordinates of the 16 train stations. It contains information on the top 10 most common venues for each of the train station locations, within a 300m radius.

Methodology

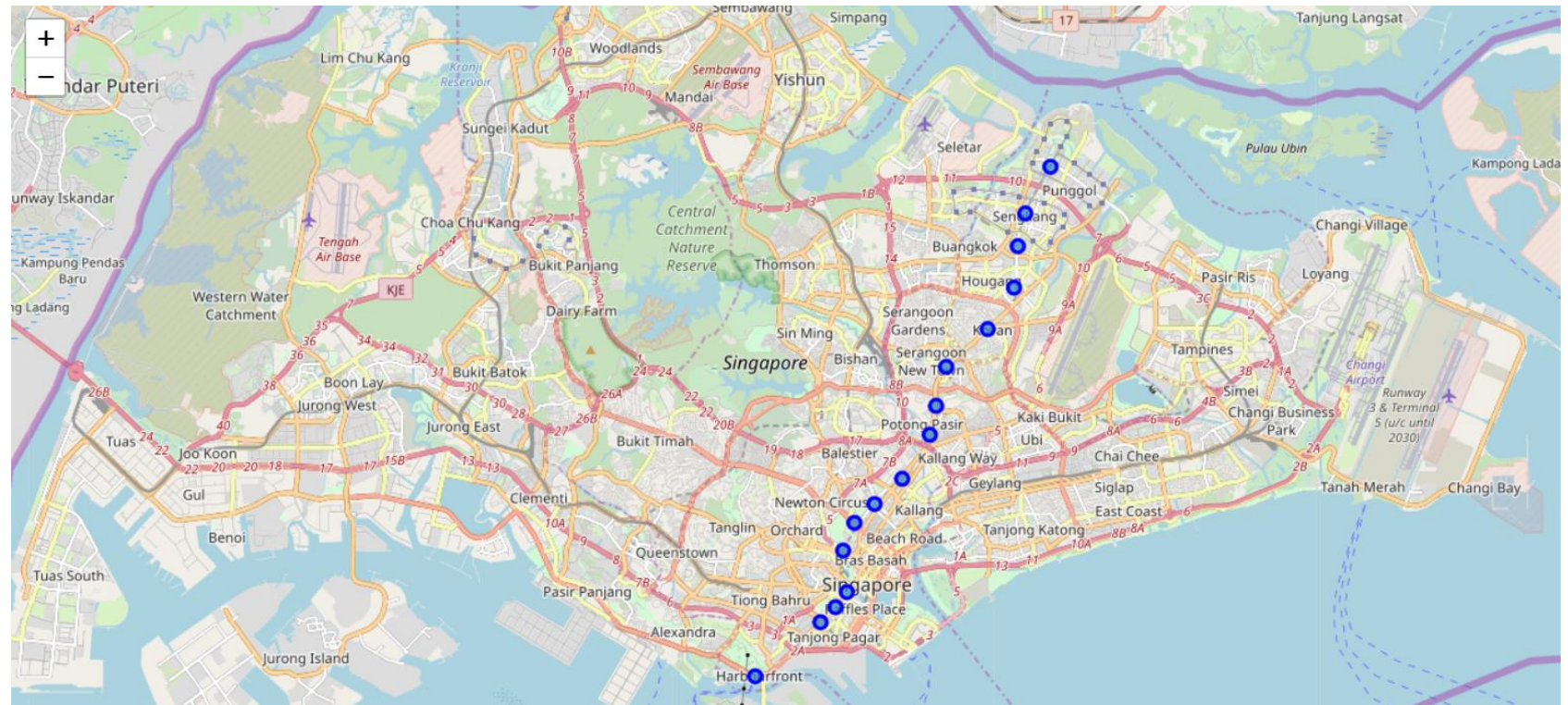
- Convert the train station dataset from a CSV file to a Pandas dataframe
- Retain only the necessary columns, i.e. station name, latitude and longitude (Note: Station name is a combination of the name of the train station as well as its station code)
- Overlay the train station locations on a map of Singapore
- Obtain FourSquare data (nearby venues - within a 300m radius) for these 16 train stations
- Determine the top ten most common venues for each of these train stations
- Evaluate what the most suitable locations would be for setting up a new cafe by identifying the train stations where cafes or coffee shops already exist, and ones where there are none
- Determine what clusters exist among the train stations using the k-means clustering method and visualise the result on a map

Visualising the train station dataset

	FullName	Latitude	Longitude
17	BOON KENG MRT STATION NE9	1.319396	103.861679
23	BUANGKOK MRT STATION NE15	1.382878	103.893104
37	CHINATOWN MRT STATION NE4	1.284360	103.843424
44	CLARKE QUAY MRT STATION NE5	1.288386	103.846552
53	DHOBY GHAUT MRT STATION NE6	1.299705	103.845485
63	FARRER PARK MRT STATION NE8	1.312360	103.854172
69	HARBOURFRONT MRT STATION NE1	1.265473	103.821446
74	HOUGANG MRT STATION NE14	1.371292	103.892364
89	KOVAN MRT STATION NE13	1.360179	103.885050
97	LITTLE INDIA MRT STATION NE7	1.307198	103.848581
117	OUTRAM PARK MRT STATION NE3	1.280225	103.839486
127	POTONG PASIR MRT STATION NE10	1.331380	103.869046
131	PUNGGOL MRT STATION NE17	1.404547	103.902052
147	SENGKANG MRT STATION NE16	1.391695	103.895466
150	SERANGOON MRT STATION NE12	1.349708	103.873564
184	WOODLEIGH MRT STATION NE11	1.339190	103.870808



Location of the 16 train stations on the North-East line



Extracting the top 10 most common venues for each station

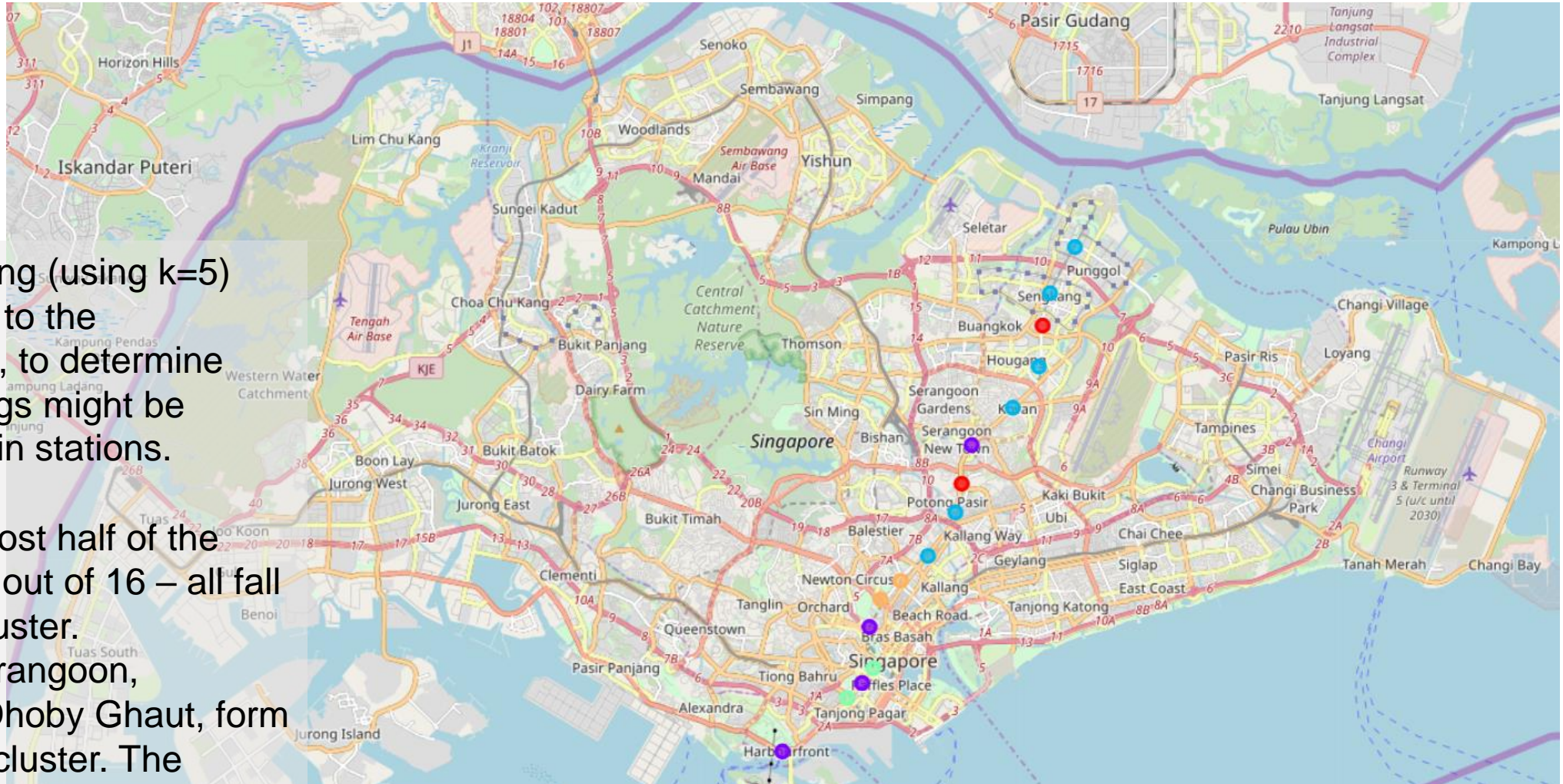
- Kovan, Hougang and Sengkang train stations have the both cafes and coffee shops mentioned amongst their top 10 most common venues
- Half of the train stations (i.e., 8 out of 16) do not have either venue type within their top 10
- Of these 8 stations, Boon Keng, Buangkok and Chinatown stand out as stations that have higher number of eateries (i.e., 5 to 9).

Station	Station	Cluster Labels	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	10th Most Common Venue	Number of Coffee Places	Number of Eateries
89	KOVAN MRT STATION NE13	2	Noodle House	Café	Asian Restaurant	Chinese Restaurant	Coffee Shop	Building	Soccer Field	Snack Place	Food Court	Sandwich Place	2	8
74	HOUGANG MRT STATION NE14	2	Fast Food Restaurant	Coffee Shop	Chinese Restaurant	Food Court	Kitchen Supply Store	Supermarket	Sandwich Place	Fish & Chips Shop	Shopping Mall	Café	2	7
147	SENGKANG MRT STATION NE16	2	Fast Food Restaurant	Café	Food Court	Sandwich Place	Sushi Restaurant	Restaurant	English Restaurant	Pastry Shop	Coffee Shop	Shopping Mall	2	7
117	OUTRAM PARK MRT STATION NE3	3	Japanese Restaurant	Australian Restaurant	Chinese Restaurant	Café	Italian Restaurant	Tapas Restaurant	Dessert Shop	Seafood Restaurant	Cocktail Bar	Gastropub	1	9
97	LITTLE INDIA MRT STATION NE7	4	Indian Restaurant	Beer Garden	Coffee Shop	Neighborhood	Cantonese Restaurant	Museum	Noodle House	Breakfast Spot	North Indian Restaurant	Thai Restaurant	1	7
131	PUNGGOL MRT STATION NE17	2	Chinese Restaurant	Multiplex	Breakfast Spot	Fast Food Restaurant	Coffee Shop	Sandwich Place	Clothing Store	Seafood Restaurant	Shopping Mall	Food Court	1	7
63	FARRER PARK MRT STATION NE8	4	Café	Indian Restaurant	Hotel	Chinese Restaurant	Sporting Goods Shop	Pakistani Restaurant	Bakery	Dumpling Restaurant	Bubble Tea Shop	Vegetarian / Vegan Restaurant	1	6
127	POTONG PASIR MRT STATION NE10	2	Noodle House	Fast Food Restaurant	Food Court	Train	Bus Station	Flower Shop	Snack Place	Coffee Shop	Convenience Store	Market	1	5
17	BOON KENG MRT STATION NE9	2	Chinese Restaurant	Noodle House	Asian Restaurant	Food Court	Bakery	Dessert Shop	Soup Place	Fast Food Restaurant	Vegetarian / Vegan Restaurant	Sandwich Place	0	9
23	BUANGKOK MRT STATION NE15	0	Wings Joint	Seafood Restaurant	Fast Food Restaurant	Basketball Court	Metro Station	Bus Station	Yoga Studio	Gastropub	English Restaurant	Fish & Chips Shop	0	6
37	CHINATOWN MRT STATION NE4	1	Chinese Restaurant	Hostel	Beer Garden	Spa	Seafood Restaurant	Food Court	Dessert Shop	Dumpling Restaurant	Pharmacy	Bed & Breakfast	0	5
44	CLARKE QUAY MRT STATION NE5	3	Seafood Restaurant	Hotel	Yoga Studio	Lounge	Spanish Restaurant	Cocktail Bar	Bubble Tea Shop	Mexican Restaurant	River	Bistro	0	4
69	HARBOURFRONT MRT STATION NE1	1	Clothing Store	Yoga Studio	Swiss Restaurant	German Restaurant	Movie Theater	Multiplex	Frozen Yogurt Shop	Food Court	Plaza	Department Store	0	4
150	SERANGOON MRT STATION NE12	1	Ice Cream Shop	Clothing Store	Chinese Restaurant	Supermarket	Bubble Tea Shop	Pet Store	Dessert Shop	Cosmetics Shop	Scenic Lookout	Shopping Mall	0	4
53	DHOBY GHAUT MRT STATION NE6	1	Yoga Studio	Soup Place	Park	Movie Theater	Miscellaneous Shop	Indonesian Restaurant	Hotpot Restaurant	Gym	Gift Shop	General Entertainment	0	3
184	WOODLEIGH MRT STATION NE11	0	Bus Station	Nightclub	Bus Line	Theater	Metro Station	Gym	Track	Frozen Yogurt Shop	Fruit & Vegetable Store	Yoga Studio	0	1

Clusters amongst the North-East line train stations

K-means clustering (using $k=5$) was also applied to the FourSquare data, to determine what the groupings might be across the 16 train stations.

Interestingly, almost half of the train stations – 6 out of 16 – all fall under a single cluster. Harbourfront, Serangoon, Chinatown and Dhoby Ghaut, form the next biggest cluster. The remainder all form clusters of two stations.



Discussion and Conclusion

- The table reflects that the majority of train stations on the North-East line already have at least one cafe or coffee shop. In particular, **Kovan, Hougang and Sengkang stations have both cafes and coffee shops** amongst their top 10 most common venues. Hence, it might be best to avoid setting up a new cafe at these stations, where there are established cafes.
- **Boon Keng, Buangkok and Chinatown** stand out as train stations with a good number of eateries (between 5 to 9) amongst its top 10 most common venues, where there is no mention of cafes or coffee shops. These are likely to be the most promising areas for setting up a new cafe.

