

Singapore MRT Station Clustering

SG : MRT Station Clustering

Singapore is one of the place that travelers would like to visit once
SG has many attractive places to see like Singapore flyer, USS, Marina Bay and many



**Moreover,
SG's MRT could link to
every parts of SG**



SG : MRT Station Clustering



Sometime
Travelers
have time left
after visit
famous places



**And they don't know
where should they go**

SG : MRT Station Clustering

Objective

To be able to guide travelers which
MRT station they should visit
based on their preference

SG : MRT Station Clustering

Data Acquisition

1. Kaggle's SG MRT coordinate, contributed by Lee Yu Xuam
*<https://www.kaggle.com/yxlee245/singapore-train-station-coordinates>
2. FoursquareAPI on 750 meter radius from MRT coordinate

SG : MRT Station Clustering

Data Cleaning

1. Remove the last 's' letter from every categories
2. Make all categories lowercased to prevent error from case sensitive

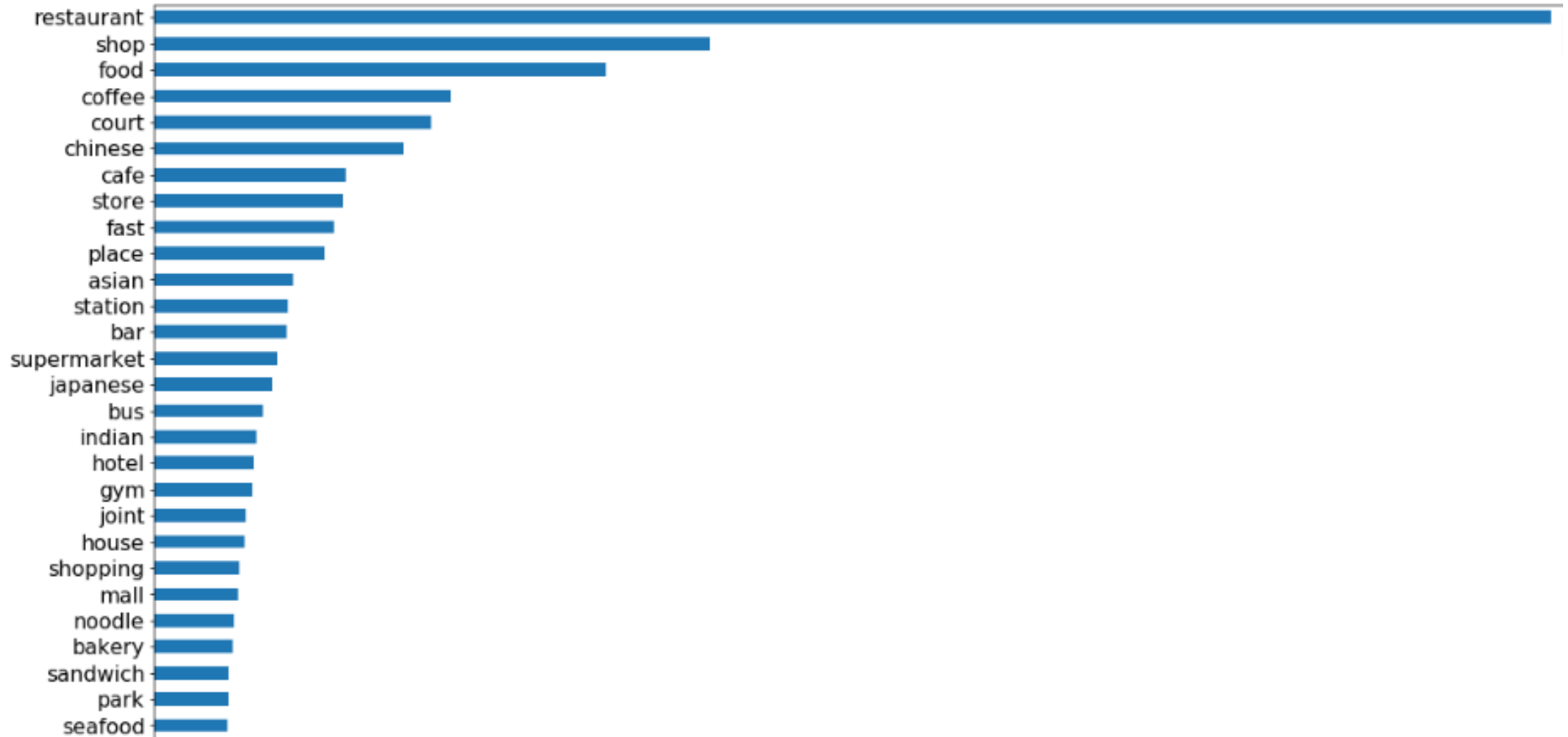
Feature Extraction

Split every word from categories name, grouping and cleaning
To add more attribute into the categories

index	cat_name	airport	alley	american	apartment	arcade	area	arena	aristocrat	...	warehouse	water	waterfall	waterfront	whisky	wine
airport	2	2	0	0	0	0	0	0	0	...	0	0	0	0	0	0
airport lounge	2	2	0	0	0	0	0	0	0	...	0	0	0	0	0	0
american restaurant	10	0	0	10	0	0	0	0	0	...	0	0	0	0	0	0
arcade	2	0	0	0	0	2	0	0	0	...	0	0	0	0	0	0

SG : MRT Station Clustering

Gain information from Attributes density



SG : MRT Station Clustering

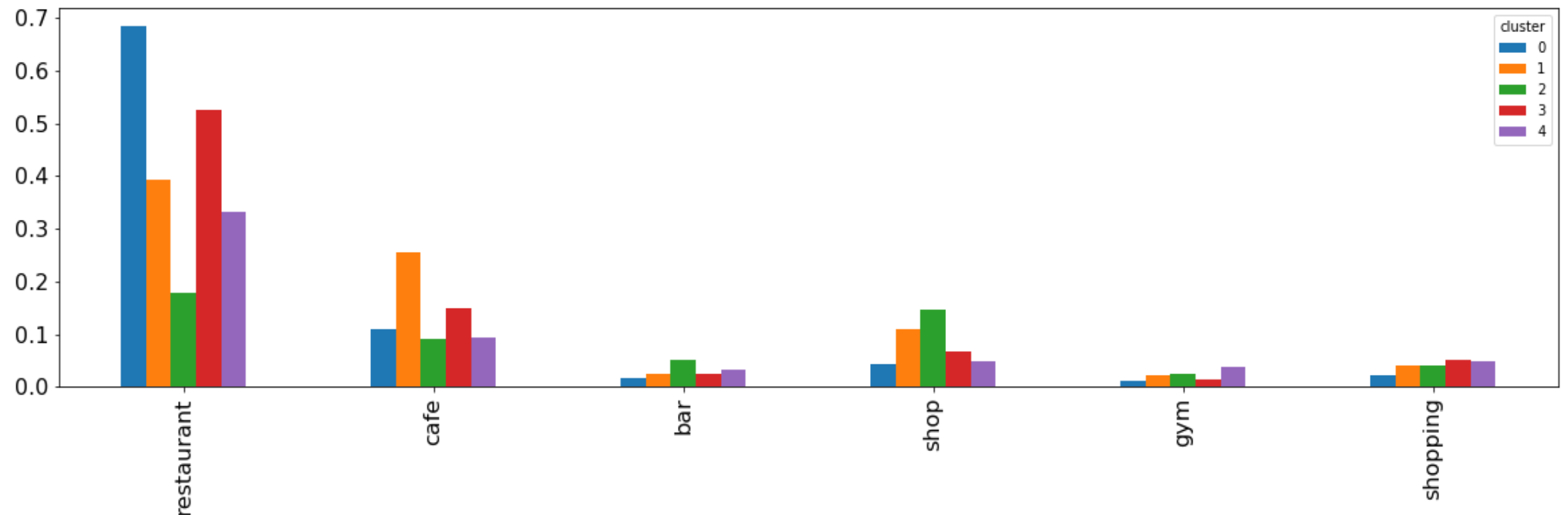
Grouping Attributes

1. Restaurant
2. Café

3. Bar
4. Shop

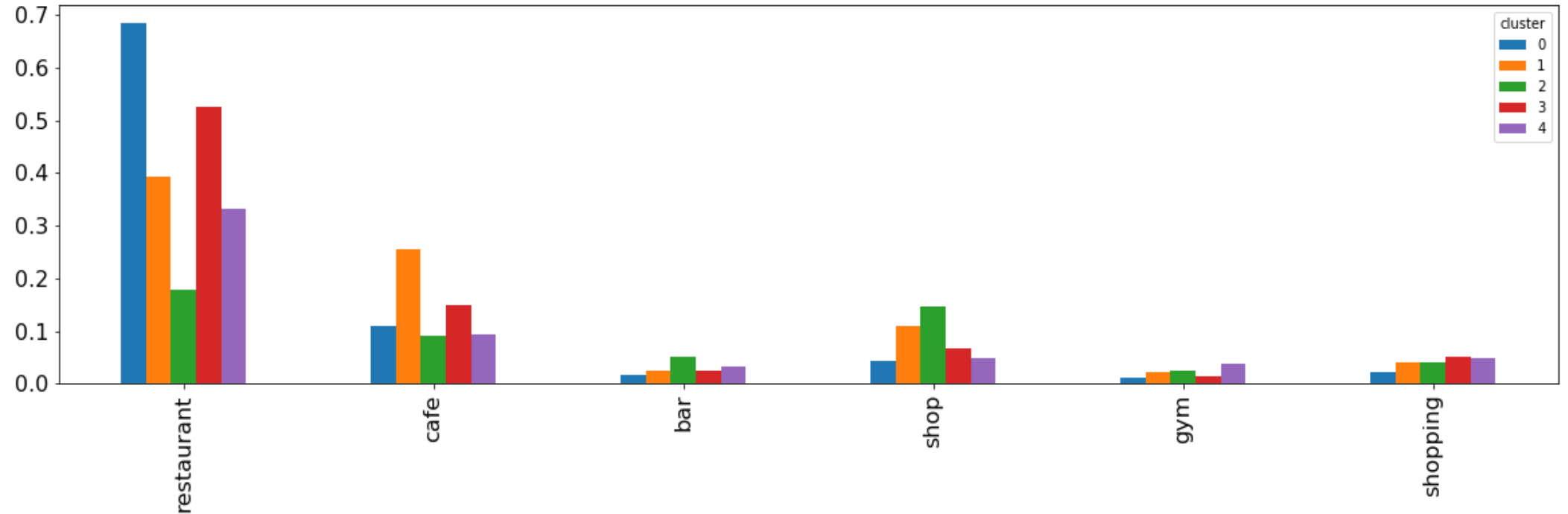
5. Gym
6. Shopping

Make the Clustering into 5 Cluster



SG : MRT Station Clustering

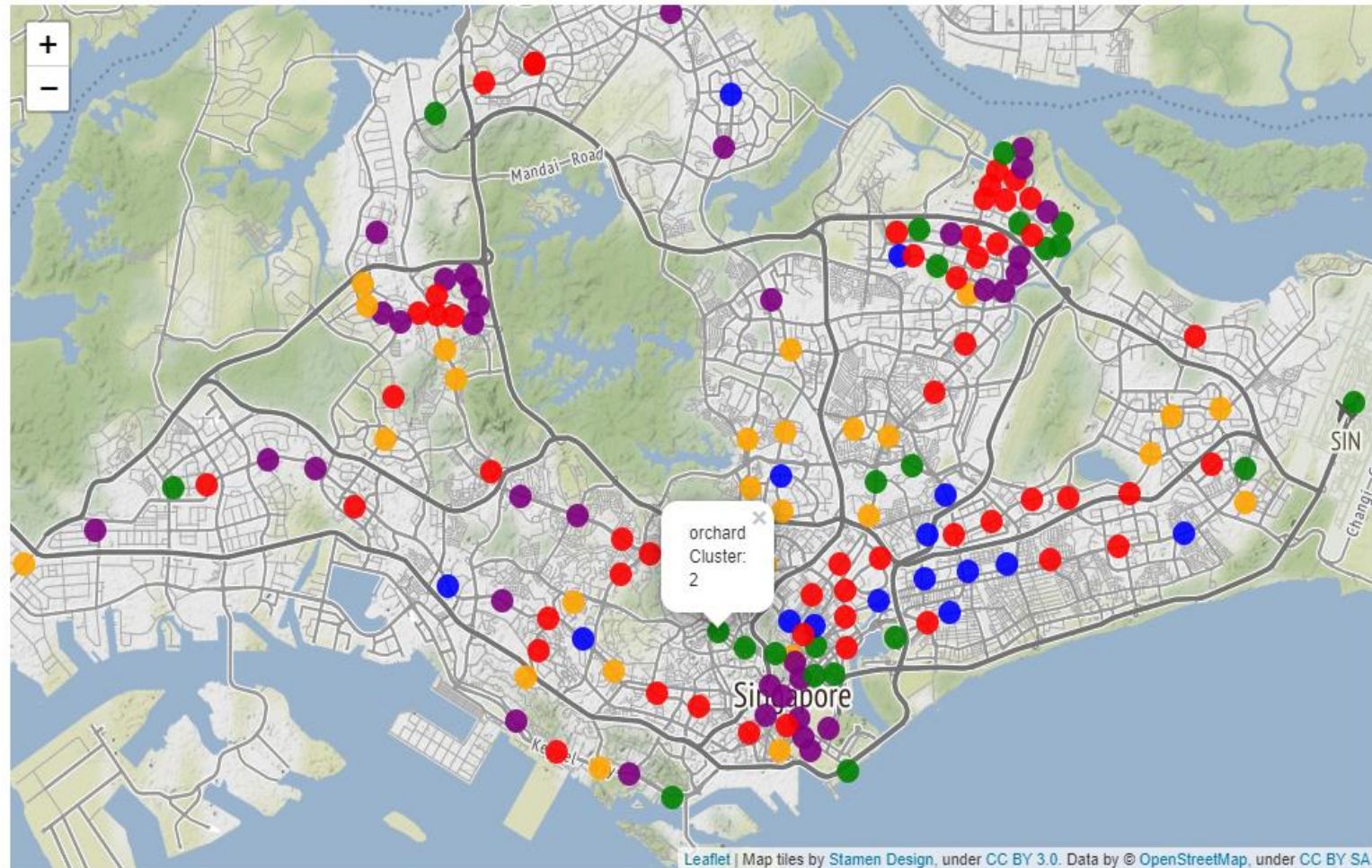
Under stand more about clusters



- Cluster 0 : Focusing on Restaurant
- Cluster 1 & 3: Focusing on Café
- Cluster 2 : Focusing on shop and bar
- Cluster 4 : Mixture of every categories

SG : MRT Station Clustering

Plotting cluster results into the map



- Cluster 0 : Focusing on Restaurant
- Cluster 1 : Focusing on Café
- Cluster 2 : Focusing on shop and bar
- Cluster 3 : Focusing on Café
- Cluster 4 : Mixture of every categories

Orchard, the popular shopping place and Harbourfront were clustered as cluster 2

And Most of station in the center of SG are cluster 4

Which are making sense to me

SG : MRT Station Clustering

Recommend Model from Clustering

From the model I build, I could receive preference from users and make model recommend cluster type for them

<div><div>Cafe lover</div><div><div>1ML_Recommend()</div><div>for 0 - 5, what do you think about : restaurant 1 for 0 - 5, what do you think about : cafe 5 for 0 - 5, what do you think about : bar 3 for 0 - 5, what do you think about : shop 0 for 0 - 5, what do you think about : gym 0 for 0 - 5, what do you think about : shopping 0 array([1])</div></div></div>	<div><div>Food lover</div><div><div>1ML_Recommend()</div><div>for 0 - 5, what do you think about : restaurant 5 for 0 - 5, what do you think about : cafe 0 for 0 - 5, what do you think about : bar 0 for 0 - 5, what do you think about : shop 0 for 0 - 5, what do you think about : gym 0 for 0 - 5, what do you think about : shopping 0 array([0])</div></div></div>	<div><div>Shopping Girls</div><div><div>1ML_Recommend()</div><div>for 0 - 5, what do you think about : restaurant 1 for 0 - 5, what do you think about : cafe 3 for 0 - 5, what do you think about : bar 2 for 0 - 5, what do you think about : shop 5 for 0 - 5, what do you think about : gym 0 for 0 - 5, what do you think about : shopping 5 array([2])</div></div></div>
C1 → Focusing on Café	C0 → Focusing on Restaurant	C2 → Focusing on shop and bar

SG : MRT Station Clustering

Conclusion

From this project, I found that

1. We could divide SG's MRT Station into 5 clusters based on attributes of places around them
2. We could recommend which station the user should go during their time based on their preference

Further more, I got a little of concept of how google or youtube recommend something to us.

Thank you!!

