

# THE RESTAURANT SCENE IN SINGAPORE

## **1. INTRODUCTION**

### **1.1. Background**

Singapore is a popular business and tourist center in southeast Asia. It also serves as one of the busiest transit points for people traveling across Asia. Singapore boasts of rich diversity in its population not only across various age groups but also across different ethnic groups. Singapore is well known for its hawker food centers and its vibrant restaurant scene which is constantly evolving. Several dining restaurants and chefs have been recognized at an international level (<https://www.stb.gov.sg/content/stb/en/industries/dining-and-retail.html>). The diverse culture along with the dining experience also helps Singapore in the tourism industry which is one of the major contributors to the country's economy.

Singapore is the second densest country (by population) in the world (<https://www.singstat.gov.sg/find-data/search-by-theme/population/population-and-population-structure/latest-data>). For sustainable living and to aid urban planning, the Singapore government divided the land into 55 planning areas (PA) across 5 regions. The census data for each of these PA is provided by the Statistics Department of Singapore.

### **1.2. Problem**

This project is aimed at entrepreneurs to help them understand the restaurant scene in Singapore.

- a) What are the different types of restaurants across different planning areas?
- b) Where should an entrepreneur open an Indian restaurant based on the ethnic population?

## **2. DATA**

### **2.1. Data Sources**

The data used in this project is collected across various publicly available data platforms.

1. The geographical details of the various Planning Areas (PA) are obtained from [https://en.wikipedia.org/wiki/Planning\\_Areas\\_of\\_Singapore](https://en.wikipedia.org/wiki/Planning_Areas_of_Singapore). The table columns used are 'Planning Area', 'Region' and 'Area (km<sup>2</sup>)'.
2. The latest population data for different PA is collected from <https://www.singstat.gov.sg/find-data/search-by-theme/population/population-and-population-structure/latest-data> (updated 24/9/2020). The total population is obtained from the 'Residents by age group and type of dwelling' dataset

3. The resident population by planning area and ethnic group is available only for 2015 ([https://data.gov.sg/dataset/resident-population-by-planning-area-subzone-ethnic-group-and-sex-2015?resource\\_id=d683afc9-d1e6-45a1-8d51-073710b7daca](https://data.gov.sg/dataset/resident-population-by-planning-area-subzone-ethnic-group-and-sex-2015?resource_id=d683afc9-d1e6-45a1-8d51-073710b7daca)) . The fraction of each ethnic group in each PA is assumed to be the same for 2020.
4. The geolocations of the various PA are obtained using Nominatim which uses OpenStreetMap data to find locations based on Name and Address (<https://nominatim.org/>).
5. The details of the types of restaurants and other venue categories for different PA are collected using the FourSquare API with a search radius personalized on the PA area.

The datasets after cleaning will be used for

## 2.2.Data cleaning and Wrangling

### 2.2.1. 2020 Population dataset

The latest population dataset is in the form of a csv file. Using Pandas, the file was read and stored as a dataframe.

	PA		SZ	AG	Sex		TOD	Pop	Time
0	Ang Mo Kio	Ang Mo Kio Town Centre	0_to_4	Males		HDB 1- and 2-Room Flats	0	2011	
1	Ang Mo Kio	Ang Mo Kio Town Centre	0_to_4	Males		HDB 3-Room Flats	10	2011	
2	Ang Mo Kio	Ang Mo Kio Town Centre	0_to_4	Males		HDB 4-Room Flats	30	2011	
3	Ang Mo Kio	Ang Mo Kio Town Centre	0_to_4	Males		HDB 5-Room and Executive Flats	50	2011	
4	Ang Mo Kio	Ang Mo Kio Town Centre	0_to_4	Males		HUDC Flats (excluding those privatised)	0	2011	

For the current problem, we require only the columns ‘PA’, ‘Pop’ for ‘Time’=2020. The data is then grouped by PA and the corresponding total population was calculated for each PA. Some of the planning areas like ‘Central Water Catchment’ do not have any population. Such rows were removed from the final dataset. This resulted in the number of Planning areas with population reducing from 55 to 42.

### 2.2.2. Planning Area/Regions and Area

The names of the different planning areas (PA) and the respective area in km<sup>2</sup> were read from [https://en.wikipedia.org/wiki/Planning\\_Areas\\_of\\_Singapore](https://en.wikipedia.org/wiki/Planning_Areas_of_Singapore) using Pandas.

	Name (English)	Malay	Chinese	Pinyin	Tamil	Region	Area (km2)	Population[7]	Density (/km2)
0	Ang Mo Kio	NaN	宏茂桥	Hóng mào qiáo	ஆங் மோ கியோ	North-East	13.94	163950	13400
1	Bedok	*	勿洛	Wù luò	பிடோக்	East	21.69	279380	13000
2	Bishan	NaN	碧山	Bì shān	பீஷான்	Central	7.62	88010	12000
3	Boon Lay	NaN	文礼	Wén lǐ	பூன் லே	West	8.23	30	3.6
4	Bukit Batok	*	武吉巴督	Wǔjí bā dū	புக்கிட் பாததோக்	West	11.13	153740	14000

The non-English columns were removed and only the Name, Region and Area columns were used. The Name column was renamed as ‘Planning Area’ and is used as the index of the resulting dataframe.

	Region	Area (km2)
<b>Planning Area</b>		
<b>Ang Mo Kio</b>	North-East	13.94
<b>Bedok</b>	East	21.69
<b>Bishan</b>	Central	7.62
<b>Boon Lay</b>	West	8.23
<b>Bukit Batok</b>	West	11.13

The above dataframe was merged with the population dataset from section 2.2.1. for common planning areas.

	Pop	Region	Area (km2)
<b>PA</b>			
<b>Ang Mo Kio</b>	162670	North-East	13.94
<b>Bedok</b>	277720	East	21.69
<b>Bishan</b>	87560	Central	7.62
<b>Bukit Batok</b>	158510	West	11.13
<b>Bukit Merah</b>	151700	Central	14.34

### 2.2.3. Ethnic group dataset

The resident population for different PA by ethnic groups is available as a csv file for the year 2015.

	year	level_1	level_2	level_3	value
<b>0</b>	2015	Total	Total	Total	3902690
<b>1</b>	2015	Total	Total	Ang Mo Kio- Total	174770
<b>2</b>	2015	Total	Total	Bedok- Total	289750
<b>3</b>	2015	Total	Total	Bishan- Total	90700
<b>4</b>	2015	Total	Total	Boon Lay- Total	30

The ‘level\_2’ column provides the total population and the population of males and females for each ethnic group. Only the rows containing ‘level\_2’= ‘Total’ are used. The ‘level\_1’ contains the ethnic groups classified as ‘Total’, ‘Chinese’, ‘Malay’, ‘Indians’ and ‘Others’. The columns ‘Year’ and ‘level\_2’ are dropped.

	level_1	value
level_3		
Ang Mo Kio	Total	174770
Bedok	Total	289750
Bishan	Total	90700
Boon Lay	Total	30
Bukit Batok	Total	139270

The idea is to obtain the fraction of each ethnic group for different PA. This is obtained by first separating the dataframe by ethnic groups.

	typet	popt	typet	popc	typem	popm	typei	popi	typeo	popo
level_3										
Ang Mo Kio	Total	174770	Chinese	143290	Malays	13060	Indians	14150	Others	4270
Bedok	Total	289750	Chinese	208880	Malays	43980	Indians	25110	Others	11780
Bishan	Total	90700	Chinese	77220	Malays	3760	Indians	7000	Others	2720
Boon Lay	Total	30	0	0	0	0	0	0	0	0
Bukit Batok	Total	139270	Chinese	102080	Malays	19570	Indians	13730	Others	3900

Some of the PA in 2015 showed 0 population for different ethnic groups. This could be possibly due to error in the census collection. Such rows are removed and ethnic fractions are calculated. The total number of PA with valid data is now 41. The ethnic fraction dataframe is now merged with the population data obtained in section 2.2.2. Two columns are added showing the ethnic group with the largest and 2<sup>nd</sup> largest fraction for each PA. The index is reset and the column PA is renamed as Planning Area.

	Planning Area	Pop	Region	Area (km2)	Chinese	Malay	Indians	Others	Max	2nd
0	Ang Mo Kio	162670	North-East	13.94	0.819878	0.074727	0.080964	0.024432	Chinese	Indians
1	Bedok	277720	East	21.69	0.720897	0.151786	0.086661	0.040656	Chinese	Malay
2	Bishan	87560	Central	7.62	0.851378	0.041455	0.077178	0.029989	Chinese	Indians
3	Bukit Batok	158510	West	11.13	0.732965	0.140518	0.098585	0.028003	Chinese	Malay
4	Bukit Merah	151700	Central	14.34	0.786768	0.085986	0.097023	0.030223	Chinese	Indians

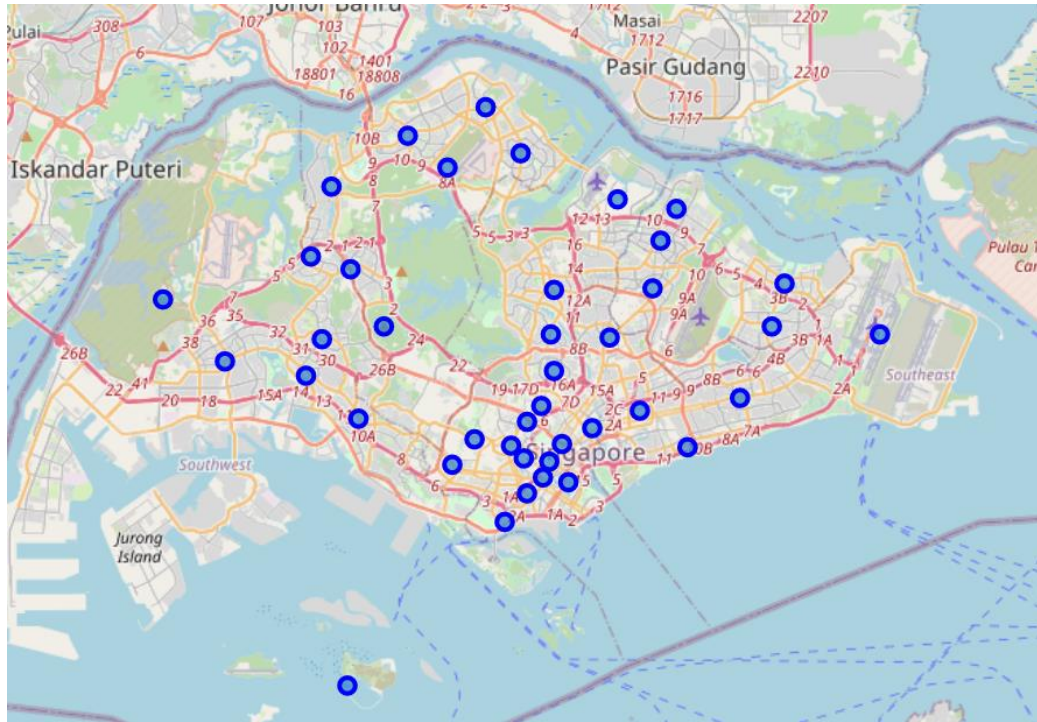
The main assumption in this study is that ethnic fractions for each PA is the same for 2020 as it was in 2015 regardless of the change in population.

#### 2.2.4. Geolocation of each PA

The coordinates of Singapore are obtained using Nominatim which uses OpenStreetMap data. The resulting latitude and longitude (1.357107, 103.8194992) is used as the starting

location to obtain the coordinates of each planning area (PA). The latitudes and longitudes are then stored as separate columns in the dataframe and is plotted using the Folium library.

	Planning Area	Pop	Region	Area (km2)	Chinese	Malay	Indians	Others	Max	2nd	Latitude	Longitude
0	Ang Mo Kio	162670	North-East	13.94	0.819878	0.074727	0.080964	0.024432	Chinese	Indians	1.370080	103.849523
1	Bedok	277720	East	21.69	0.720897	0.151786	0.086661	0.040656	Chinese	Malay	1.323976	103.930216
2	Bishan	87560	Central	7.62	0.851378	0.041455	0.077178	0.029989	Chinese	Indians	1.350986	103.848255
3	Bukit Batok	158510	West	11.13	0.732965	0.140518	0.098585	0.028003	Chinese	Malay	1.349057	103.749591
4	Bukit Merah	151700	Central	14.34	0.786768	0.085986	0.097023	0.030223	Chinese	Indians	1.270439	103.828318



### 2.2.5. Nearby Venues using FourSquare API

The FourSquare API is used to obtain the nearby venues for each PA based on a search radius. The search radius in ‘m’ for each PA is calculated using the area of the PA as follows

$$Radius = 1000 \sqrt{\frac{Area}{\pi}}$$

Based on the above radius, the FourSquare API reports the nearby venues along with the venue latitude and longitude and venue category.

	Planning Area	Latitude	Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Ang Mo Kio	1.37008	103.849523	Bishan - Ang Mo Kio Park	1.362219	103.846250	Park
1	Ang Mo Kio	1.37008	103.849523	Aramsa ~ The Garden Spa	1.362292	103.847602	Spa
2	Ang Mo Kio	1.37008	103.849523	Old Chang Kee	1.369094	103.848389	Snack Place
3	Ang Mo Kio	1.37008	103.849523	FairPrice Xtra	1.369279	103.848886	Supermarket
4	Ang Mo Kio	1.37008	103.849523	Bangkok Street Mookata	1.365688	103.853186	BBQ Joint

Only restaurants and related categories like bar/pub/food courts are used for further analysis. The analysis is performed for each PA and is then studied along with the population dataset obtained in section 2.2.4.

### 3. METHODOLOGY

The FourSquare API is used to obtain 100 nearby venues for each PA out of which there are 304 unique venue categories (from the dataframe obtained in section 2.2.5). I have used one hot encoding, a method to assign 1 or 0 to each venue category into a new column. This is done using the `get_dummies` function. For the present study, only restaurant categories including food courts and bars are required. 81 restaurant categories are available across Singapore.

	Planning Area	American Restaurant	Asian Restaurant	Australian Restaurant	BBQ Joint	Bar	Beach Bar	Beer Bar	Burger Joint	Cantonese Restaurant	Chinese Aristocrat Restaurant	Chinese Restaurant	Cocktail Bar	Comfort Food Restaurant	Cuban Restaurant
0	Ang Mo Kio	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	Ang Mo Kio	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	Ang Mo Kio	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	Ang Mo Kio	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	Ang Mo Kio	0	0	0	0	0	0	0	1	0	0	0	0	0	0

The categories are then grouped by each PA and the mean is calculated for each type of restaurant.

	Planning Area	American Restaurant	Asian Restaurant	Australian Restaurant	BBQ Joint	Bar	Beach Bar	Beer Bar	Burger Joint	Cantonese Restaurant	Chinese Aristocrat Restaurant	Chinese Restaurant	Cocktail Bar	Comfort Food Restaurant	Cuban Restaurant
0	Ang Mo Kio	0.00	0.04	0.01	0.03	0.02	0.0	0.0	0.01	0.0	0.0	0.11	0.00	0.00	0.0
1	Bedok	0.02	0.06	0.00	0.00	0.00	0.0	0.0	0.00	0.0	0.0	0.14	0.00	0.01	0.0
2	Bishan	0.00	0.04	0.00	0.03	0.02	0.0	0.0	0.01	0.0	0.0	0.10	0.00	0.00	0.0
3	Bukit Batok	0.01	0.01	0.00	0.01	0.01	0.0	0.0	0.00	0.0	0.0	0.12	0.00	0.00	0.0
4	Bukit Merah	0.00	0.01	0.01	0.00	0.01	0.0	0.0	0.00	0.0	0.0	0.03	0.01	0.00	0.0

It was observed that the largest ethnic group for all the PA was Chinese. Therefore, I have used 2<sup>nd</sup> largest group as one of the categories for each PA. Indians – 0, Malay – 1 and Others – 3.

K-means clustering is an unsupervised algorithm, which divides data into K clusters. Data within clusters have similar properties. This algorithm minimizes the intra-cluster distance and maximizes the inter-cluster distances. The PA are divided into clusters based on the types of restaurant and the 2<sup>nd</sup> largest ethnic group.

For the second part of analysis, I have collected all the PA which have venue category ‘Indian Restaurant’. The Indian population along with the Indian population density for each of these PA was calculated using the ethnic fraction from section 2.2.4. The PA

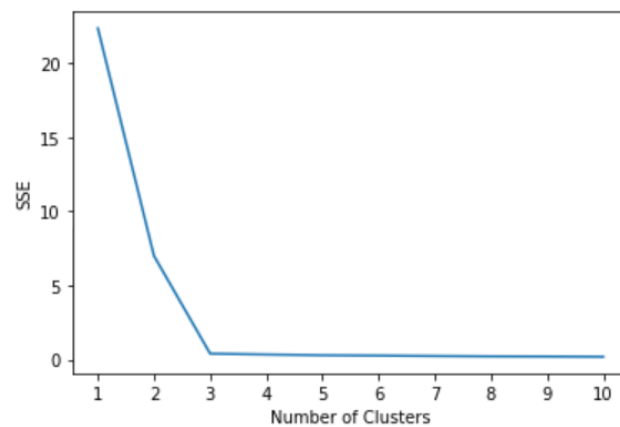


containing the 2<sup>nd</sup> largest ethnic groups as Indians and also Indian Restaurants were then collected.

	Planning Area	Indian Restaurants	Pop	Region	Area (km2)	Indians	2nd	ind_pop	ind_pd
0	Ang Mo Kio	1	162670	North-East	13.94	0.080964	Indians	13170.0	945.0
1	Bishan	1	87560	Central	7.62	0.077178	Indians	6758.0	887.0
2	Hougang	1	228130	North-East	13.93	0.081148	Indians	18512.0	1329.0
3	Kallang	4	101680	Central	9.17	0.142871	Indians	14527.0	1584.0
4	Marine Parade	4	46230	Central	6.12	0.123333	Indians	5702.0	932.0
5	Novena	13	49440	Central	8.98	0.087727	Indians	4337.0	483.0
6	Rochor	20	12730	Central	1.62	0.202879	Indians	2583.0	1594.0
7	Serangoon	1	117070	North-East	10.10	0.089003	Indians	10420.0	1032.0

#### 4. RESULTS & DISCUSSION

The optimal number of clusters (K) is obtained using the elbow method.



At K=3, there is a trade-off between error and the number of clusters. The clusters are plotted using the Folium library.



Pink – Cluster 0; Red – Cluster 1; Green – Cluster 2

The details for each cluster are as follows:

	Planning Area	Pop	Max	2nd	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	Density
0	Ang Mo Kio	162670	Chinese	Indians	Chinese Restaurant	Food Court	Noodle House	Japanese Restaurant	Asian Restaurant	11669.0
2	Bishan	87560	Chinese	Indians	Chinese Restaurant	Food Court	Asian Restaurant	Seafood Restaurant	Noodle House	11491.0
4	Bukit Merah	151700	Chinese	Indians	Japanese Restaurant	Chinese Restaurant	Korean Restaurant	Food Court	Sushi Restaurant	10579.0
12	Hougang	228130	Chinese	Indians	Chinese Restaurant	Food Court	Fast Food Restaurant	Noodle House	Asian Restaurant	16377.0
15	Kallang	101680	Chinese	Indians	Chinese Restaurant	Thai Restaurant	Restaurant	Italian Restaurant	Indian Restaurant	11088.0
16	Mandai	2060	Chinese	Indians	Food Court	Fast Food Restaurant	Malay Restaurant	Noodle House	Dim Sum Restaurant	175.0
17	Marine Parade	46230	Chinese	Indians	Noodle House	Asian Restaurant	Chinese Restaurant	Italian Restaurant	Bar	7554.0
20	Novena	49440	Chinese	Indians	Indian Restaurant	Chinese Restaurant	Food Court	Seafood Restaurant	Noodle House	5506.0
27	Rochor	12730	Chinese	Indians	Indian Restaurant	Vegetarian / Vegan Restaurant	Chinese Restaurant	Restaurant	Japanese Restaurant	7858.0
31	Serangoon	117070	Chinese	Indians	Chinese Restaurant	Noodle House	Asian Restaurant	Italian Restaurant	Korean Restaurant	11591.0
34	Sungei Kadut	680	Chinese	Indians	Asian Restaurant	Food Court	Fast Food Restaurant	Japanese Restaurant	Bar	43.0
38	Western Water Catchment	640	Chinese	Indians	Fast Food Restaurant	Chinese Restaurant	Food Court	Restaurant	Halal Restaurant	9.0

Cluster 0 consists of areas where Indians are the 2<sup>nd</sup> largest ethnic group. The most common type of restaurant is varied. While Chinese restaurants are most common venue since Chinese are the majority ethnic group, Asian restaurants are a common type here. Surprisingly, only 2 planning areas have Indian restaurants as the most common venue.

	Planning Area	Pop	Max	2nd	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	Density
1	Bedok	277720	Chinese	Malay	Chinese Restaurant	Food Court	Asian Restaurant	Seafood Restaurant	Japanese Restaurant	12804.0
3	Bukit Batok	158510	Chinese	Malay	Chinese Restaurant	Food Court	Japanese Restaurant	Fast Food Restaurant	Vegetarian / Vegan Restaurant	14242.0
5	Bukit Panjang	138790	Chinese	Malay	Fast Food Restaurant	Food Court	Asian Restaurant	Sushi Restaurant	Indonesian Restaurant	15438.0
7	Changi	1860	Chinese	Malay	Fast Food Restaurant	Yunnan Restaurant	Chinese Restaurant	Juice Bar	Latin American Restaurant	46.0
8	Choa Chu Kang	192480	Chinese	Malay	Fast Food Restaurant	Food Court	Chinese Restaurant	Italian Restaurant	Food & Drink Shop	31502.0
9	Clementi	92210	Chinese	Malay	Chinese Restaurant	Food Court	Japanese Restaurant	Indian Restaurant	Noodle House	9717.0
11	Geylang	110450	Chinese	Malay	Chinese Restaurant	Food Court	Asian Restaurant	Seafood Restaurant	BBQ Joint	11457.0
13	Jurong East	78840	Chinese	Malay	Chinese Restaurant	Food Court	Japanese Restaurant	Vegetarian / Vegan Restaurant	American Restaurant	4422.0
14	Jurong West	263050	Chinese	Malay	Fast Food Restaurant	Food Court	Asian Restaurant	Chinese Restaurant	Japanese Restaurant	17907.0
22	Outram	18330	Chinese	Malay	Chinese Restaurant	Asian Restaurant	Noodle House	Italian Restaurant	Japanese Restaurant	13380.0
23	Pasir Ris	147340	Chinese	Malay	Food Court	Fast Food Restaurant	Asian Restaurant	Seafood Restaurant	Italian Restaurant	9810.0
24	Punggol	174700	Chinese	Malay	Fast Food Restaurant	Food Court	Chinese Restaurant	Asian Restaurant	Japanese Restaurant	18704.0
25	Queenstown	96110	Chinese	Malay	Chinese Restaurant	Seafood Restaurant	Food Court	Asian Restaurant	French Restaurant	4704.0
29	Sembawang	102840	Chinese	Malay	Chinese Restaurant	Fast Food Restaurant	Asian Restaurant	Food Court	Seafood Restaurant	8334.0
30	Sengkang	249670	Chinese	Malay	Fast Food Restaurant	Food Court	Chinese Restaurant	Indonesian Restaurant	Sushi Restaurant	23576.0
35	Tampines	260380	Chinese	Malay	Fast Food Restaurant	Food Court	Thai Restaurant	Japanese Restaurant	Noodle House	12464.0
37	Toa Payoh	122220	Chinese	Malay	Chinese Restaurant	Food Court	Asian Restaurant	Seafood Restaurant	Noodle House	14960.0
39	Woodlands	255350	Chinese	Malay	Food Court	Fast Food Restaurant	Chinese Restaurant	Asian Restaurant	Indian Restaurant	18790.0
40	Yishun	221940	Chinese	Malay	Chinese Restaurant	Asian Restaurant	Food Court	Indian Restaurant	Thai Restaurant	10449.0

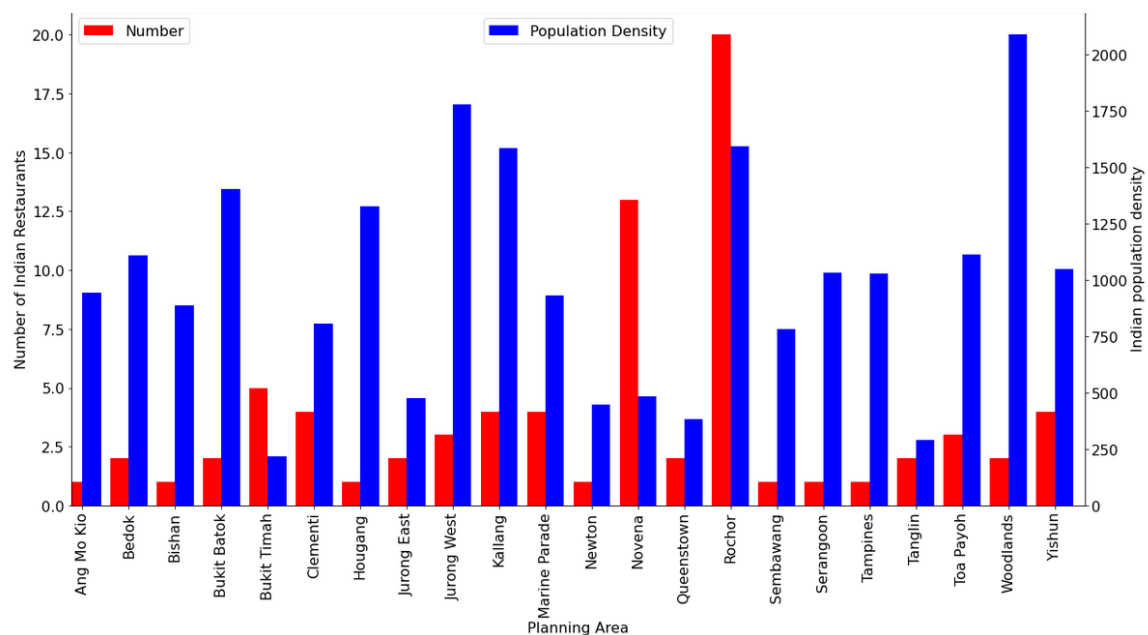


Cluster=1 are areas with Malay is the 2<sup>nd</sup> largest ethnic group. These areas contain a majority of Fast-food restaurants, food courts and Chinese restaurant.

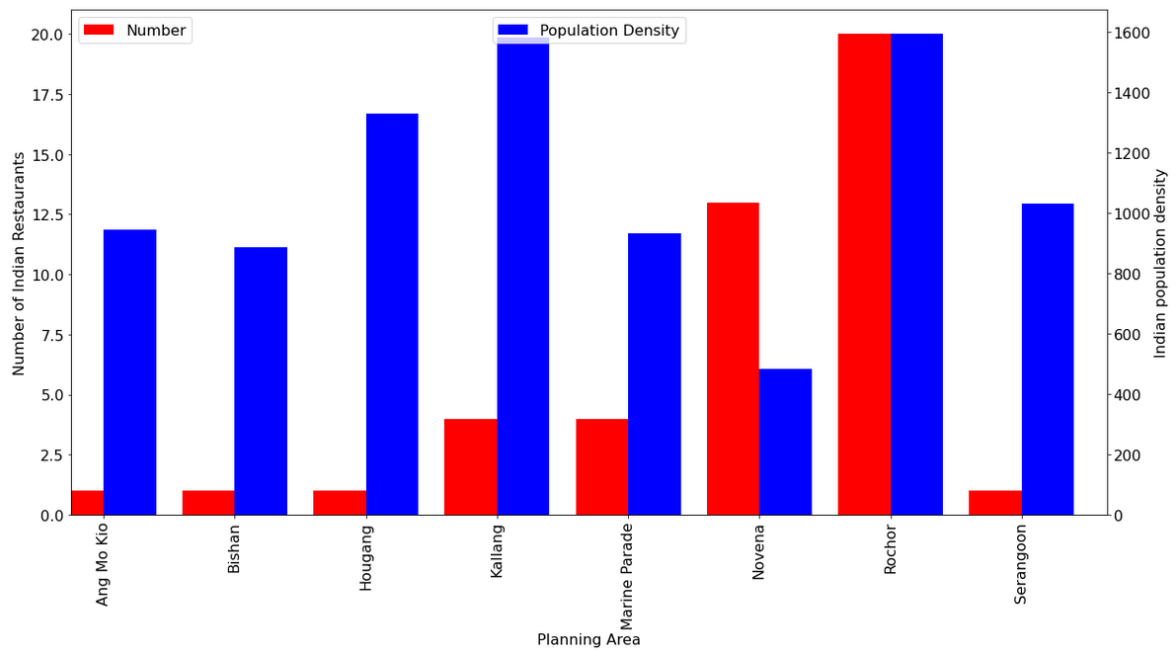
	Planning Area	Pop	Max	2nd	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	Density
6	Bukit Timah	78180	Chinese	Others	Chinese Restaurant	Korean Restaurant	Indian Restaurant	Food Court	Italian Restaurant	4460.0
10	Downtown Core	2970	Chinese	Others	Japanese Restaurant	Italian Restaurant	Cocktail Bar	Modern European Restaurant	Dim Sum Restaurant	684.0
18	Museum	500	Chinese	Others	Japanese Restaurant	Restaurant	Asian Restaurant	Whisky Bar	Bar	602.0
19	Newton	8250	Chinese	Others	Chinese Restaurant	Japanese Restaurant	Italian Restaurant	Seafood Restaurant	Asian Restaurant	3986.0
21	Orchard	810	Chinese	Others	Japanese Restaurant	Sushi Restaurant	Chinese Restaurant	Asian Restaurant	Indonesian Restaurant	844.0
26	River Valley	10090	Chinese	Others	Japanese Restaurant	Hotpot Restaurant	Chinese Restaurant	Sushi Restaurant	Dumpling Restaurant	6818.0
28	Seletar	280	Chinese	Others	Asian Restaurant	Restaurant	Southern / Soul Food Restaurant	Food Court	Bar	27.0
32	Singapore River	3100	Chinese	Others	Japanese Restaurant	Noodle House	Italian Restaurant	Bar	Food Court	3229.0
33	Southern Islands	1940	Chinese	Others	Yunnan Restaurant	Italian Restaurant	Food & Drink Shop	Food Court	Food Truck	320.0
36	Tanglin	21900	Chinese	Others	French Restaurant	Chinese Restaurant	Italian Restaurant	Modern European Restaurant	Seafood Restaurant	2870.0

The cluster consists of areas where the 2<sup>nd</sup> largest ethnic groups are neither Malay nor Indians. The restaurants in this cluster are a variety of Asian and non-Asian restaurants.

The second part is the analysis of Indian restaurants in Singapore. The following bar chart shows the planning areas that contain Indian restaurants along with the Indian population density (Indian population/Area). These areas consist of Malay or Indians as the 2<sup>nd</sup> largest ethnic group.



The data is then reduced to areas where the 2<sup>nd</sup> largest groups were Indians and also contained Indian restaurants.



From the bar plots it is seen that the best areas to open an Indian restaurant is at Kallang and Hougang which has the next best population density after Rochor which has the maximum number of Indian restaurants. However, there are areas where Indians are the 3<sup>rd</sup> largest ethnic group but still have higher Indian population density like Woodlands and Jurong west where an Indian restaurant could be successful.

This study has few uncertainties and assumptions:

- The ethnic fraction for each Planning area is available only for the year 2015 and is assumed to be the same for 2020.
- Some of the planning areas have lower population and population density. These areas are popular tourism areas in the central district area of Singapore or are present on the borders of Singapore.
- The population data is for residents and does not take into account tourists' data.

For future analysis, clustering should include tourism population along with non-residential population. Similar analysis can be made for other cities as well.

## **5. CONCLUSION**

In this capstone project, data was extracted from public platforms and using FourSquare API was used to obtain different types of restaurants for different planning areas in Singapore. The planning areas were divided into 3 clusters based on the type of restaurants and 2<sup>nd</sup> maximum ethnic group. Chinese restaurants and fast-food restaurants/food courts were part of areas that large population density and had Chinese and Malay as the largest

ethnic groups. Areas which have neither Malay or Indians as the 2<sup>nd</sup> largest ethnic groups, there is a variety of Asian and non-Asian restaurants. These areas are also popular tourist destinations. The clusters containing Indians surprisingly have Asian restaurants and lesser number of Indian restaurants. Based on the Indian population density, places like Kallang and Hougang will be a popular destination for an Indian restaurant. There are other areas where despite Indians being the 3<sup>rd</sup> largest ethnic group, the Indian population density is slightly higher than Rochor which has the maximum number of Indian restaurants like Woodlands and Jurong West where Indian restaurants could be successful.

## **6. REFERENCES**

Planning Areas of Singapore (23 March 2021):

[https://en.wikipedia.org/wiki/Planning\\_Areas\\_of\\_Singapore](https://en.wikipedia.org/wiki/Planning_Areas_of_Singapore)

Population data (23 March 2021): <https://www.singstat.gov.sg/find-data/search-by-theme/population/population-and-population-structure/latest-data>

Population by Ethnic Groups (23 March 2021): [https://data.gov.sg/dataset/resident-population-by-planning-area-subzone-ethnic-group-and-sex-2015?resource\\_id=d683afc9-d1e6-45a1-8d51-073710b7daca](https://data.gov.sg/dataset/resident-population-by-planning-area-subzone-ethnic-group-and-sex-2015?resource_id=d683afc9-d1e6-45a1-8d51-073710b7daca)

Nominatim information: <http://nominatim.org>