COURSERA Capstone Project: Diversity of Restaurants in Singapore

author: JL

Table of contents

Introduction

- Singapore has the second greatest population density in the world, and has very vibrant and diverse communities.
- As a home to a wide range of cultures, ethnicities and religions, Singapore has a rich choice of different cuisines and restaurants.
- People from same background and culture tend to gather spatially and form local communities, and it is assumed that such spatial pattern of communities can be reflected by the popularity and distribution of different types of restaurants.
- In addition, a map that presents the clusters of different types of cuisines in Singapore can be treated as a kind of food guide map for tourists and local citizens, which can be interesting despite commercial valuable

Data

- **Spatial data of Singapore**

- The Singapore subzone shapefile data can be downloaded from the following link on data.gov.sg (https://data.gov.sg/dataset/master-plan-2019-subzone-boundary-no-sea).
- There are a total of 325 Singapore subzones in the data downloaded.

	Lon	Lat	SUBZONE_NAME	SUBZONE_CODE	PLN_AREA_NAME	PLN_AREA_CODE	REGION_NAME	REGION_CODE
0	103.872352	1.288517	MARINA EAST	MESZ01	MARINA EAST	ME	CENTRAL REGION	CR
1	103.837500	1.294016	INSTITUTION HILL	RVSZ05	RIVER VALLEY	RV	CENTRAL REGION	CR
2	103.837064	1.291286	ROBERTSON QUAY	SRSZ01	SINGAPORE RIVER	SR	CENTRAL REGION	CR
3	103.698639	1.262532	JURONG ISLAND AND BUKOM	WISZ01	WESTERN ISLANDS	WI	WEST REGION	WR
4	103.846053	1.294046	FORT CANNING	MUSZ02	MUSEUM	MU	CENTRAL REGION	CR

Data

- **Food related POI (restaurants) data**

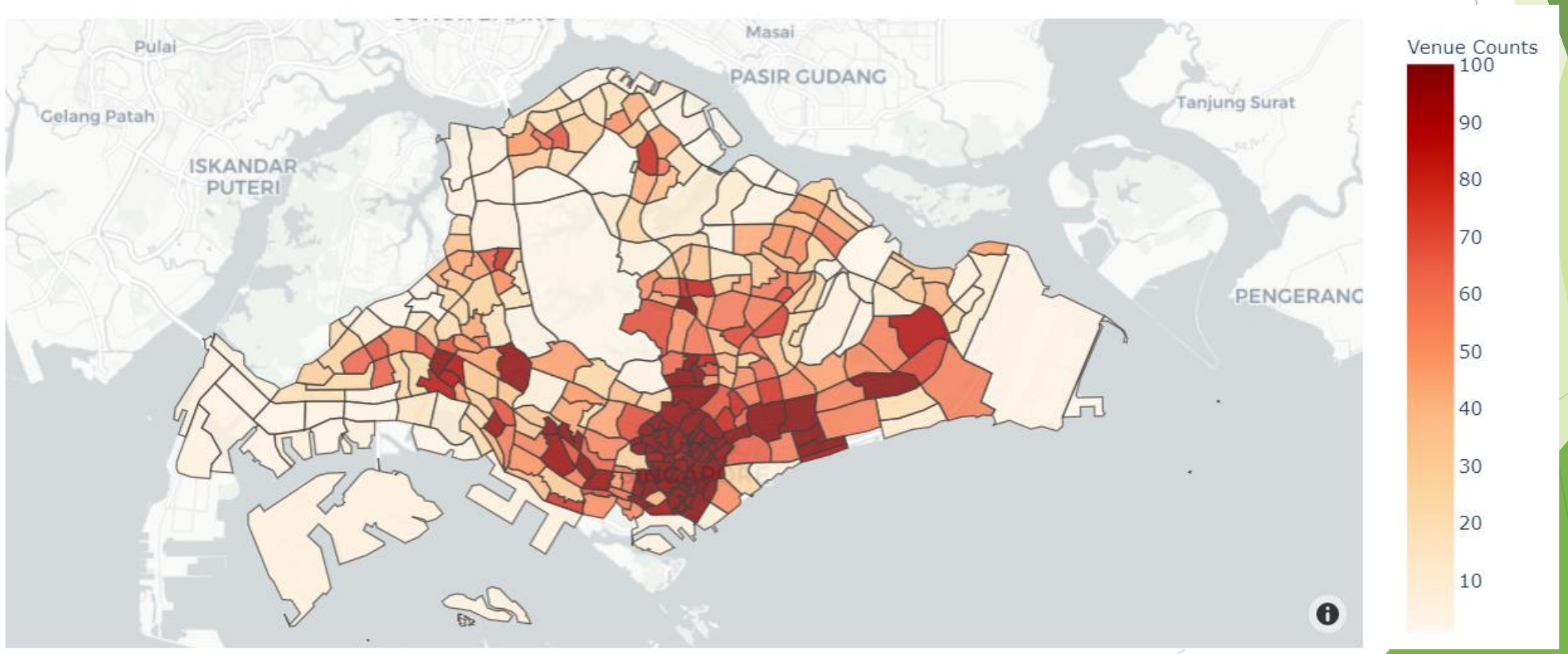
- The restaurant data can be collected from Foursquare.
- A total of 16294 venues were collected.

	name	categories	lat	lng
0	Yen Yakiniku	Japanese Restaurant	1.281074	103.845743
1	Bam! Tapas-Sake Bar	Tapas Restaurant	1.278393	103.844426
2	Tippling Club	Restaurant	1.279420	103.843848
3	PS.Cafe	Café	1.280468	103.846264
4	Fat Prince	Kebab Restaurant	1.277801	103.845202
5	Lolla	Spanish Restaurant	1.281034	103.845708
6	Park Bench Deli	Deli / Bodega	1.279872	103.847287
7	Pantler	Bakery	1.280137	103.847256
8	Maxwell Food Centre	Food Court	1.280291	103.844742
9	Dumpling Darlings	Dumpling Restaurant	1.280483	103.846942
10	Super Star K Korean BBQ	Korean Restaurant	1.278003	103.843680

Methodology

- **Exploratory Data Analysis**

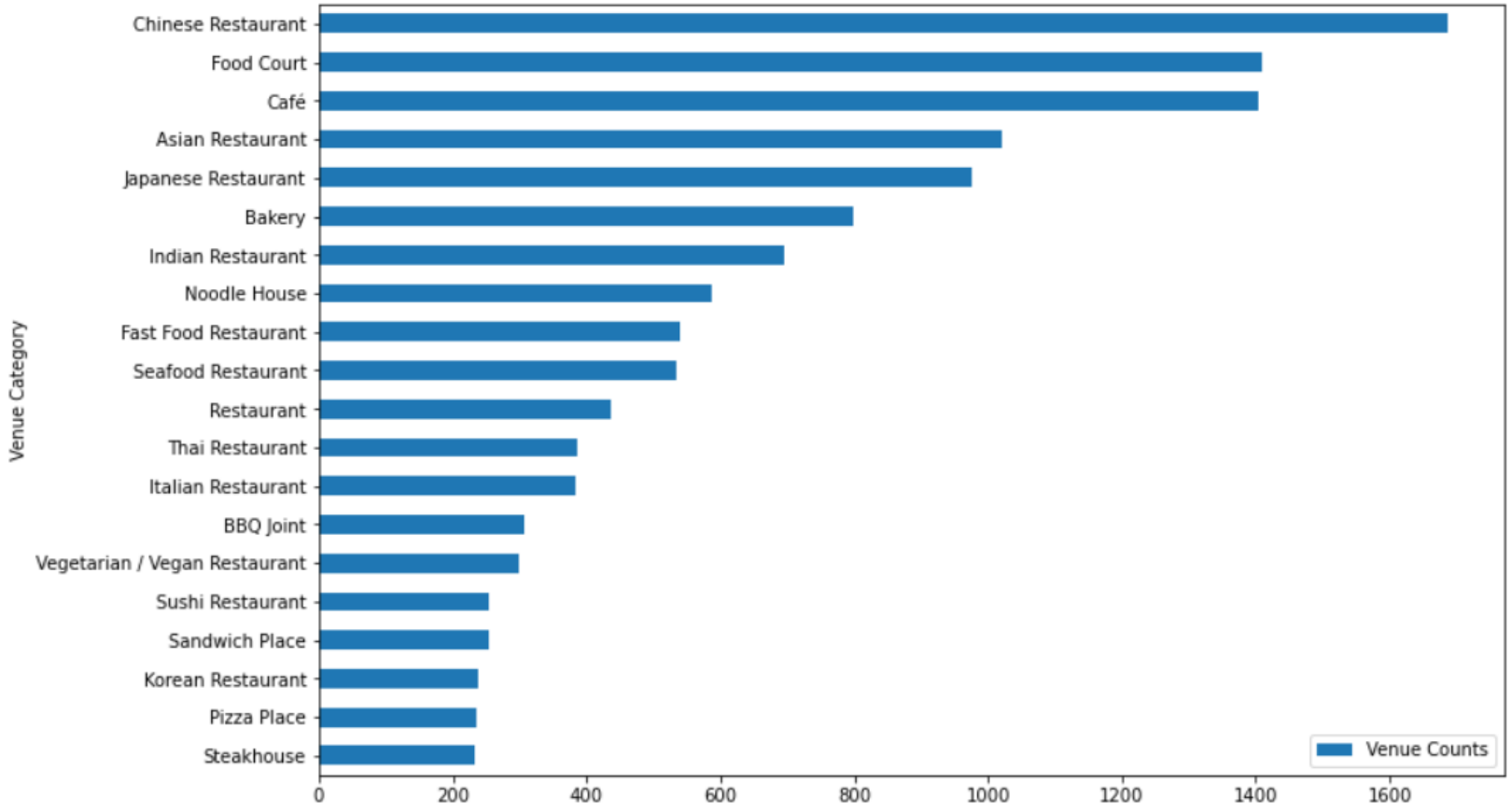
Venue counts by subzone



Methodology

- **Exploratory Data Analysis**

Venue counts by restaurant types (top 20 types displayed)



Methodology

- **Exploratory Data Analysis**

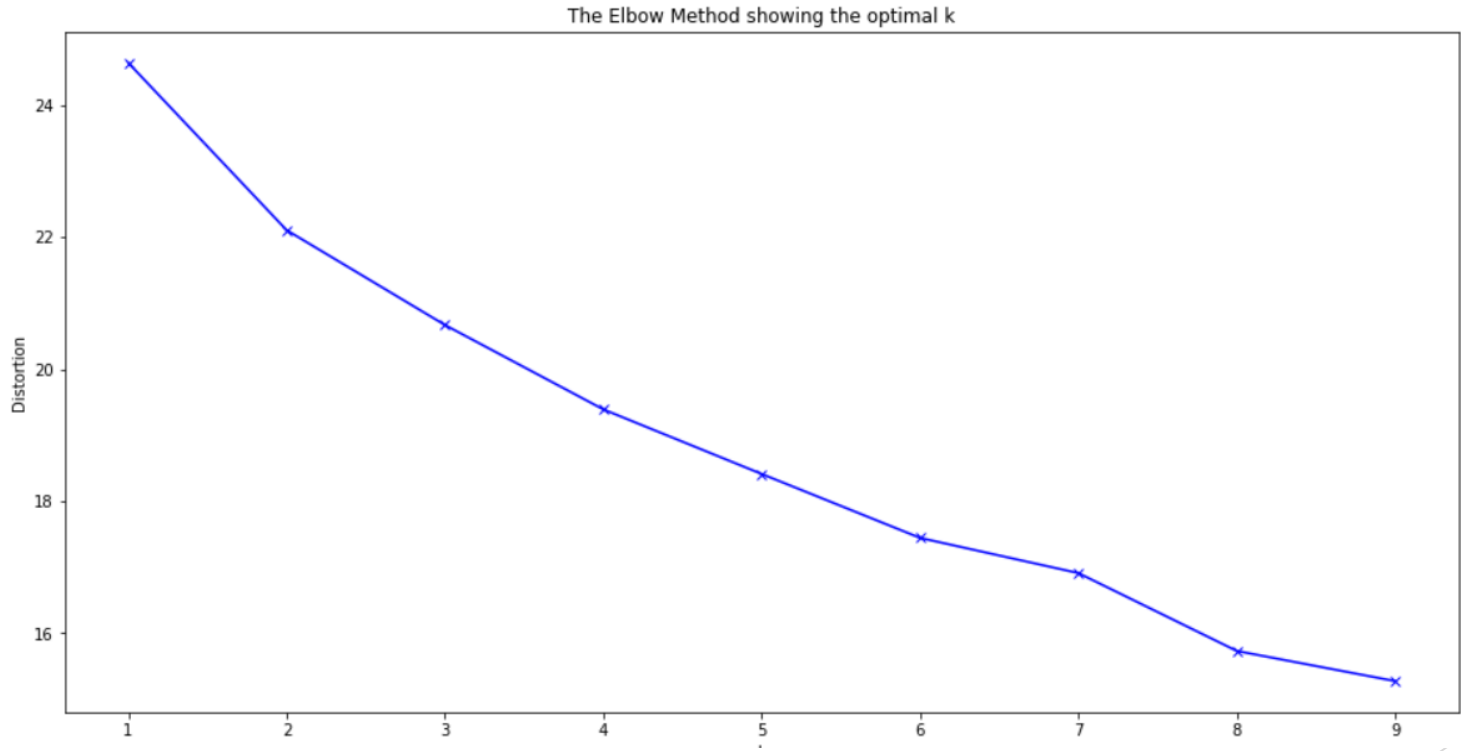
Check top 10 types for each subzone (examples displayed)

	Subzone	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
0	ADMIRALTY	Chinese Restaurant	Food Court	Fast Food Restaurant	Bakery	Soup Place	Italian Restaurant	Pizza Place	Asian Restaurant	Seafood Restaurant	Japanese Restaurant
1	AIRPORT ROAD	Food Court	Breakfast Spot	Cafeteria	Bakery	Noodle House	Chinese Restaurant	Restaurant	Asian Restaurant	Indian Restaurant	Fast Food Restaurant
2	ALEXANDRA HILL	Chinese Restaurant	Food Court	Noodle House	Café	Bakery	Asian Restaurant	BBQ Joint	Indian Restaurant	Sandwich Place	Fast Food Restaurant
3	ALEXANDRA NORTH	Chinese Restaurant	Café	Noodle House	Bakery	Food Court	Indian Restaurant	Fast Food Restaurant	Asian Restaurant	Hainan Restaurant	Bistro
4	ALJUNIED	Chinese Restaurant	Noodle House	Food Court	Dim Sum Restaurant	Seafood Restaurant	Asian Restaurant	Vegetarian / Vegan Restaurant	BBQ Joint	Café	Thai Restaurant

Methodology

- **Clustering**

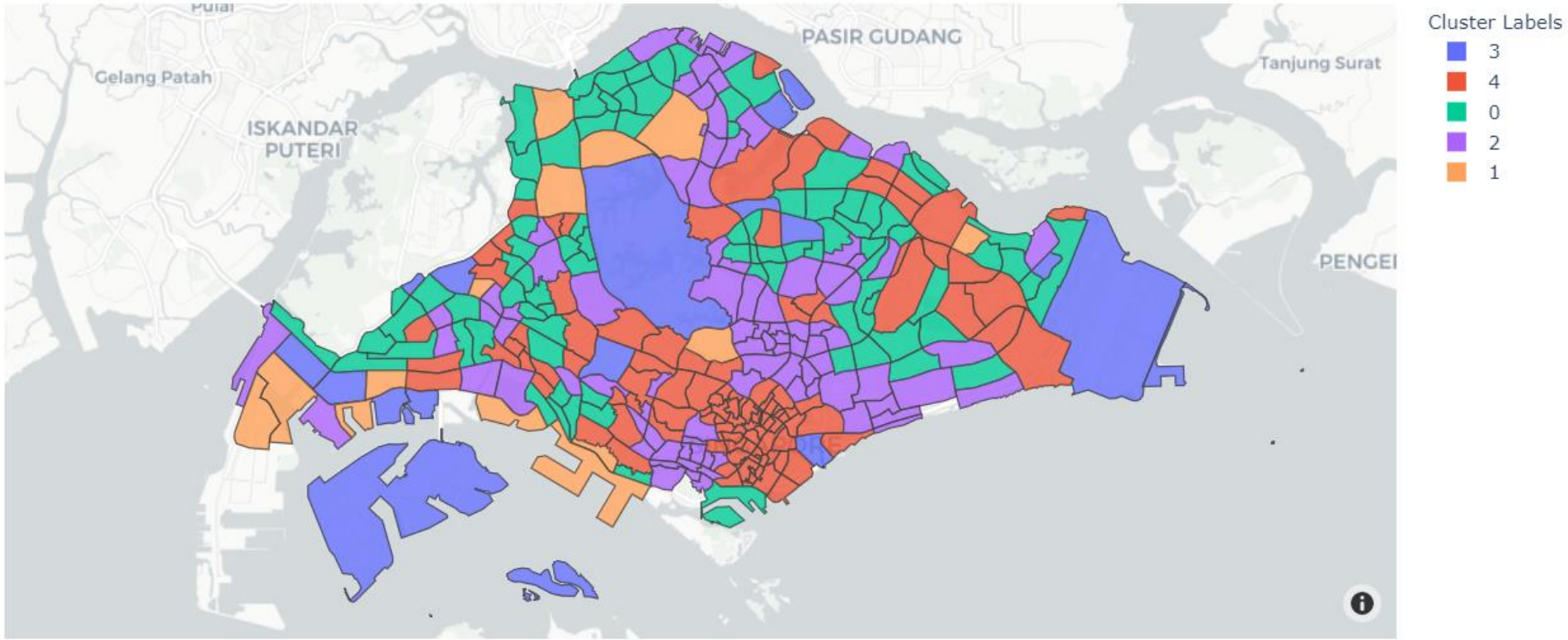
K-means clustering
Optimal k = 5 based on the Elbow Test



Results and discussions

- **Clustering**

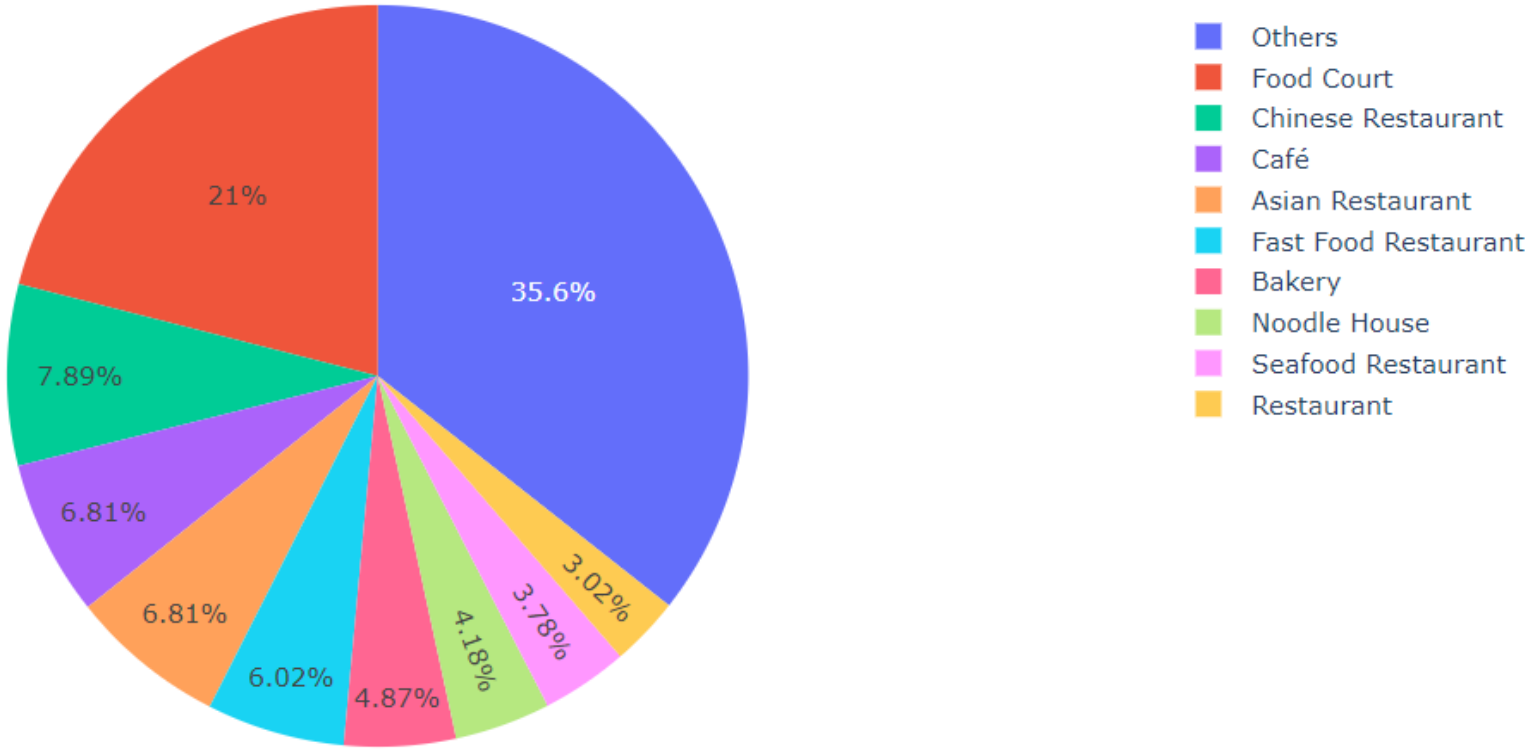
Subzone clusters based on restaurant distributions



Results and discussions

- **Cluster profile #0**

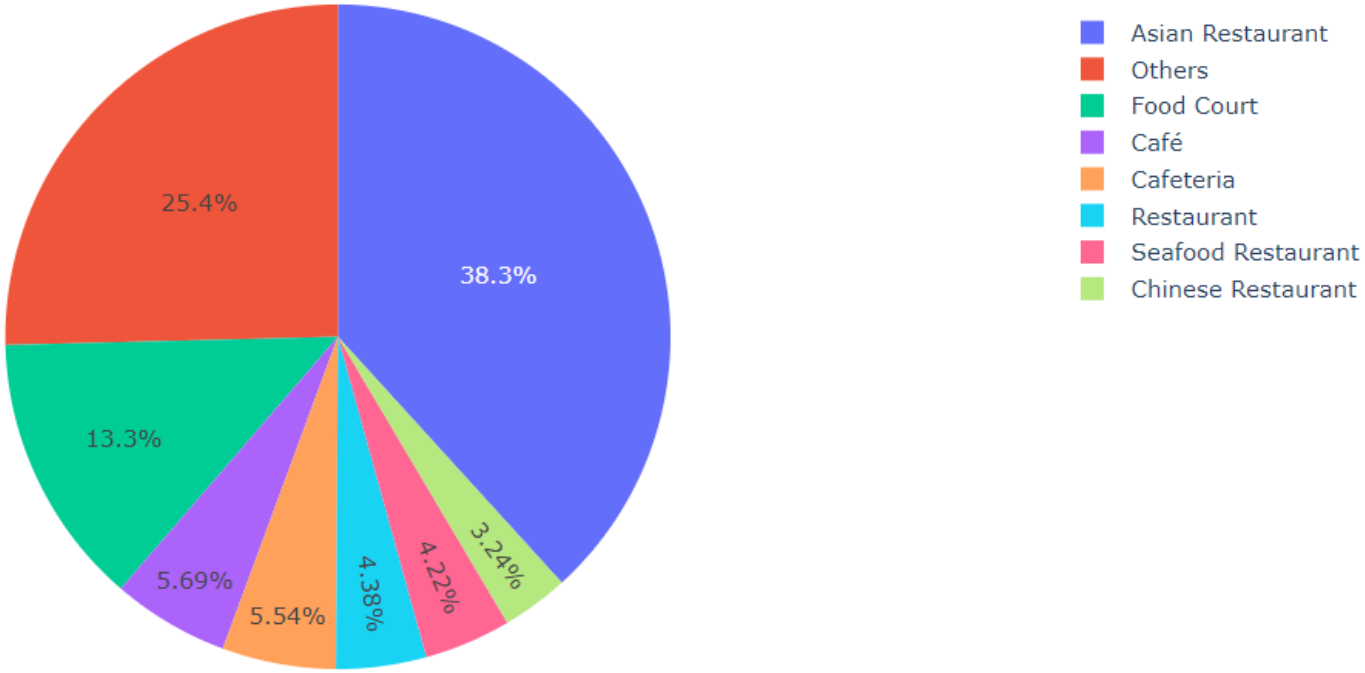
Restaurant distribution within cluster_0



Results and discussions

- **Cluster profile #1**

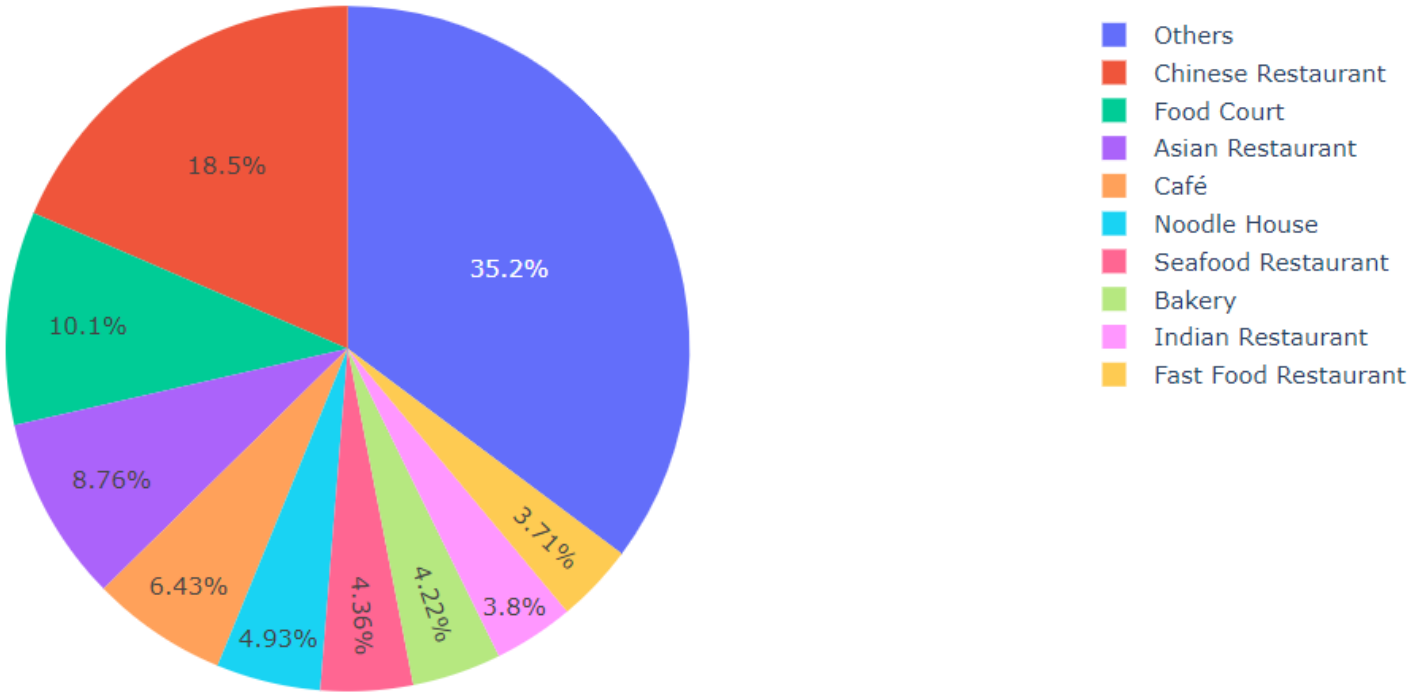
Restaurant distribution within cluster_1



Results and discussions

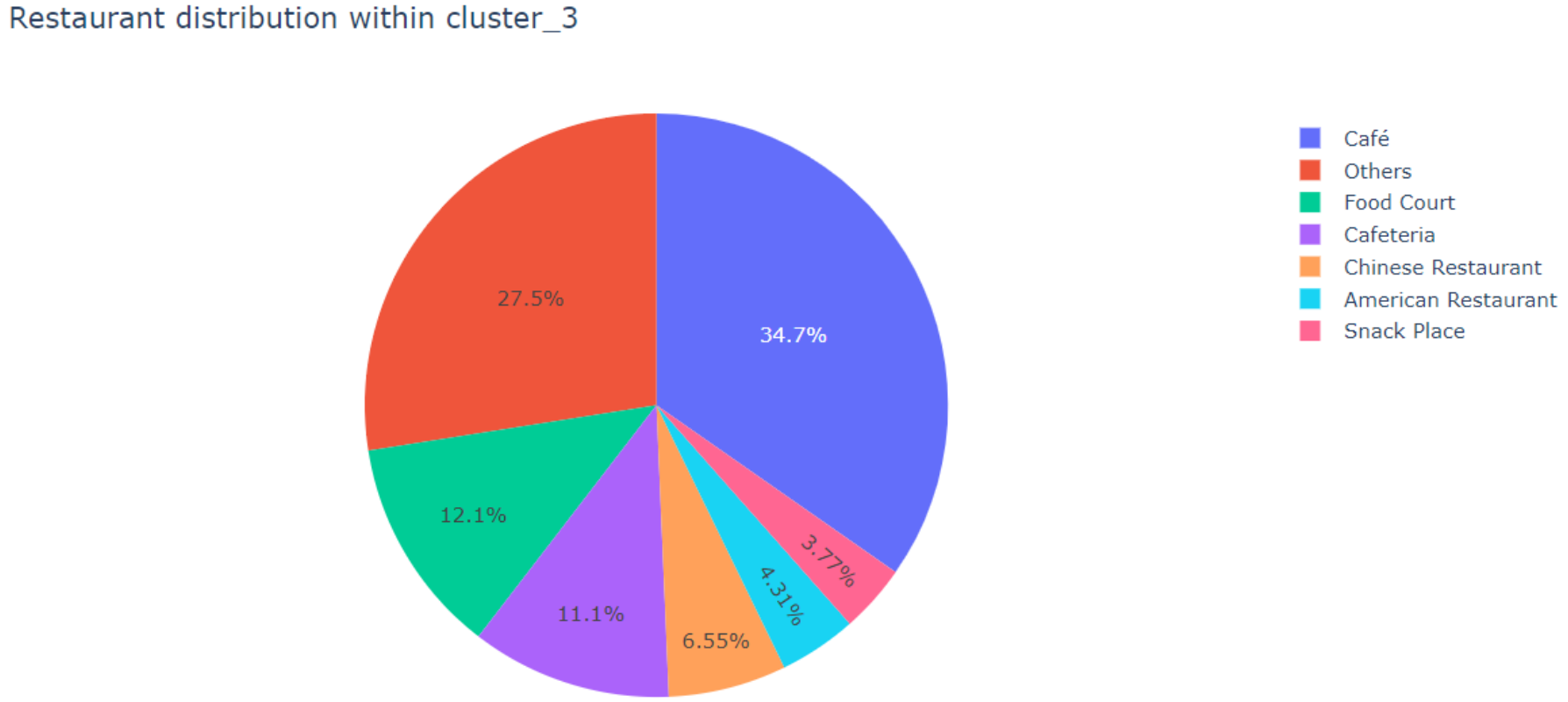
- **Cluster profile #2**

Restaurant distribution within cluster_2



Results and discussions

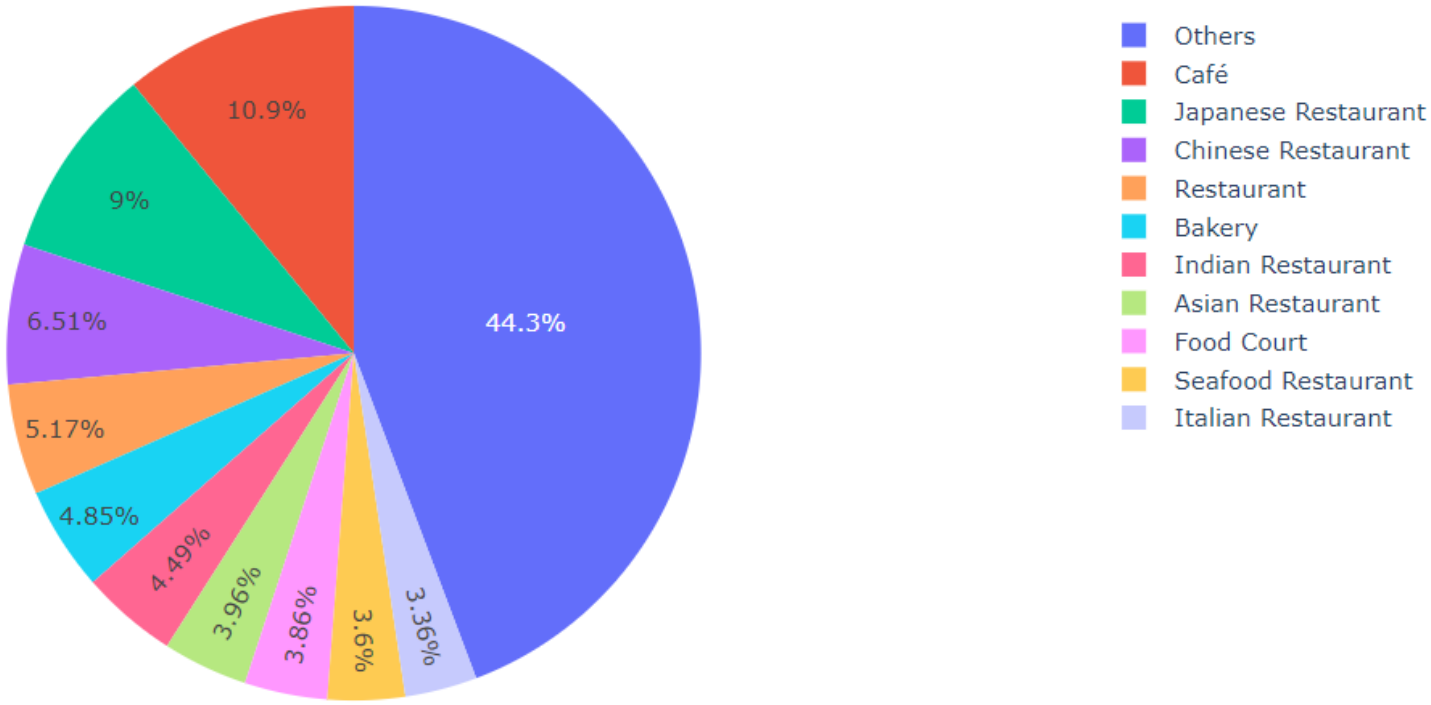
- **Cluster profile #3**



Results and discussions

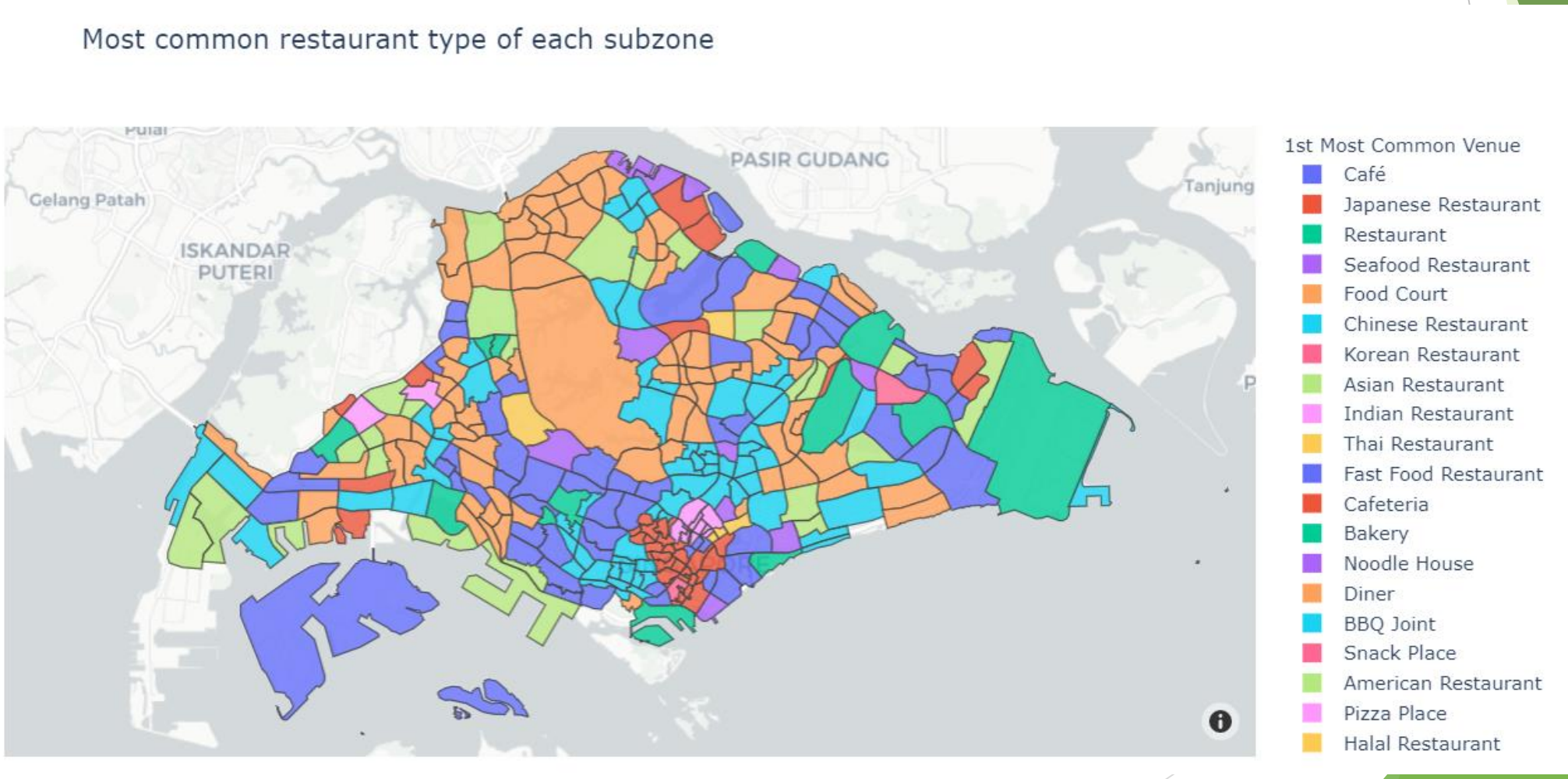
- **Cluster profile #4**

Restaurant distribution within cluster_4



Results and discussions

- **Most common restaurant type of each subzone**



Conclusions

- This project is trying to explore the diversity of restaurants in Singapore and provide an overview of distributions of different restaurant types in different subzones.
- By analysing the top 100 food-related Foursquare POIs of each subzone, we successfully identified the distribution of different cuisines in the local communities, and observed some interesting spatial patterns of the subzone clusters purely based on food preferences.
- If someone is interested in opening a restaurant in Singapore, they can use the outcome of this project for reference in popularity of the relevant restaurant types.
- This is just an practical project with very limited time. However, it can be extended and improved in many aspects.