COURSERA Capstone Project: Restaurants in Singapore

Author: JL

1. Introduction

1.1 Background

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. For example, places that have many Chinese restaurants may be the places that Chinese communities stay or visit most. 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.

1.2 Data Description

To achieve the product as described above, two major datasets are required, the spatial data of Singapore and food related Point of Interests (POI).

(1) Spatial data of Singapore

To facilitate urban planning, the Urban Redevelopment Authority (URA) divides Singapore into regions, planning areas and subzones. The Planning Regions are divided into smaller Planning Areas. Each Planning Area is further divided into smaller subzones which are usually centred around a focal point such as neighbourhood centre or activity node. There are over three hundred subzones of a total of 55 planning areas, organised into 5 regions. To achieve a more detailed investigation, this project will be conducted in the subzone level.

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.

The subzones are presented as polygons in the original shapefile. With the help of QGIS (a spatial analysis tool), we can extract the centroid of each subzone polygon (as shown in figure 1). The subzone information including latitude, longitude and name, can be exported as a csv file for the use of POI collection later, an example is given in figure 2.
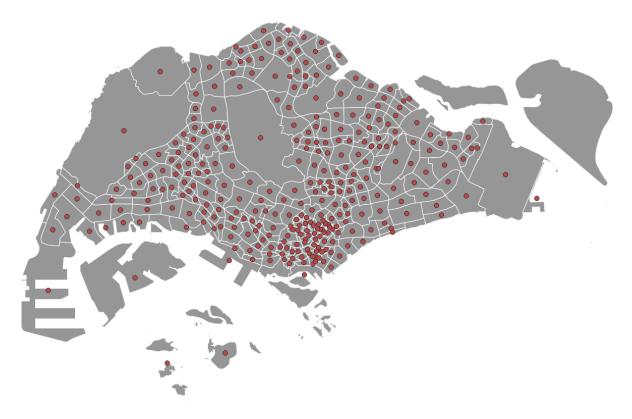
Figure 1 Singapore subzones

	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

Figure 2 Example of the SG_Subzone data

(2) Food related POI (restaurants) data

The restaurant data can be collected from Foursquare. We can search for all the POIs under the "Food" category (Foursquare categoryID is '4d4b7105d754a06374d81259') around each subzone centroid. The searching buffer is defined as 1 km to ensure a good coverage, the limit of venues returned per request was set as 100. A total of 16294 venues were collected. A sample of the collected POIs after processing is shown in figure 3. The POIs will be joint with subzones and more exploratory analysis will be conducted in the following sessions.

	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

Figure 3 data sample of collected POIs