Capstone Project

Setting up Chinese Restaurant in the Greater Johor Bahru Region

Author: Y.D. Lim

Abstract

In this capstone project, a research on a market-entry problem is presented. Johor Bahru, one of the biggest cities in Malaysia, is explored for the favorable locations of opening a Chinese restaurant. Useful data is collected and *k*-means clustering machine learning modelling is performed to sort out the similar neighborhoods. A few most promising neighborhoods and their advantages are discussed, and the potential neighborhoods to operate high-end, mid-end and family-owned (small-sized) Chinese restaurants are recommended, respectively. Lastly, it is hoped that the insights generated from the analysis will be valuable for stakeholders who are interested in opening a Chinese restaurant in the city.

Keywords: Chinese restaurant, Johor Bahru, Malaysia, *k*-means, neighborhoods

Table of Contents

Introduction

Johor Bahru is the capital city of the Johor state in Malaysia. It is the second largest city of Malaysia (following Kuala Lumpur), with a total population of about 1,061,950 and densities of 814/km^2, 1200/km^2 and 2,259/km^2 in the metro, urban and city areas, respectively [1]. The greater Johor Bahru region consists of Johor Bahru city centre, Pasir Gudang, Pulai, Gelang Patah, Kulai and Tebrau. The region is well-known of its multi-racial cultures and religions, where Malay and Chinese contribute about 48.4 % and 37.6% to the total population. In addition, Johor Bahru is vicinal to the Republic of Singapore and the Senai International Airport, which have enhanced the tourism development in the region. With the support of Iskandar Development Region project [2], the population and economic in the greater Johor Bahru area is expected to grow drastically. Hence, it is worth to explore the area for potential business opportunities.

This report will target stakeholders who are interested in opening a Chinese restaurant in the greater Johor Bahru region. The neighborhoods covered by the region will be identified and analysed. A few potential neighborhoods will be discussed so that the stakeholders may choose the best possible location to open a Chinese restaurant.

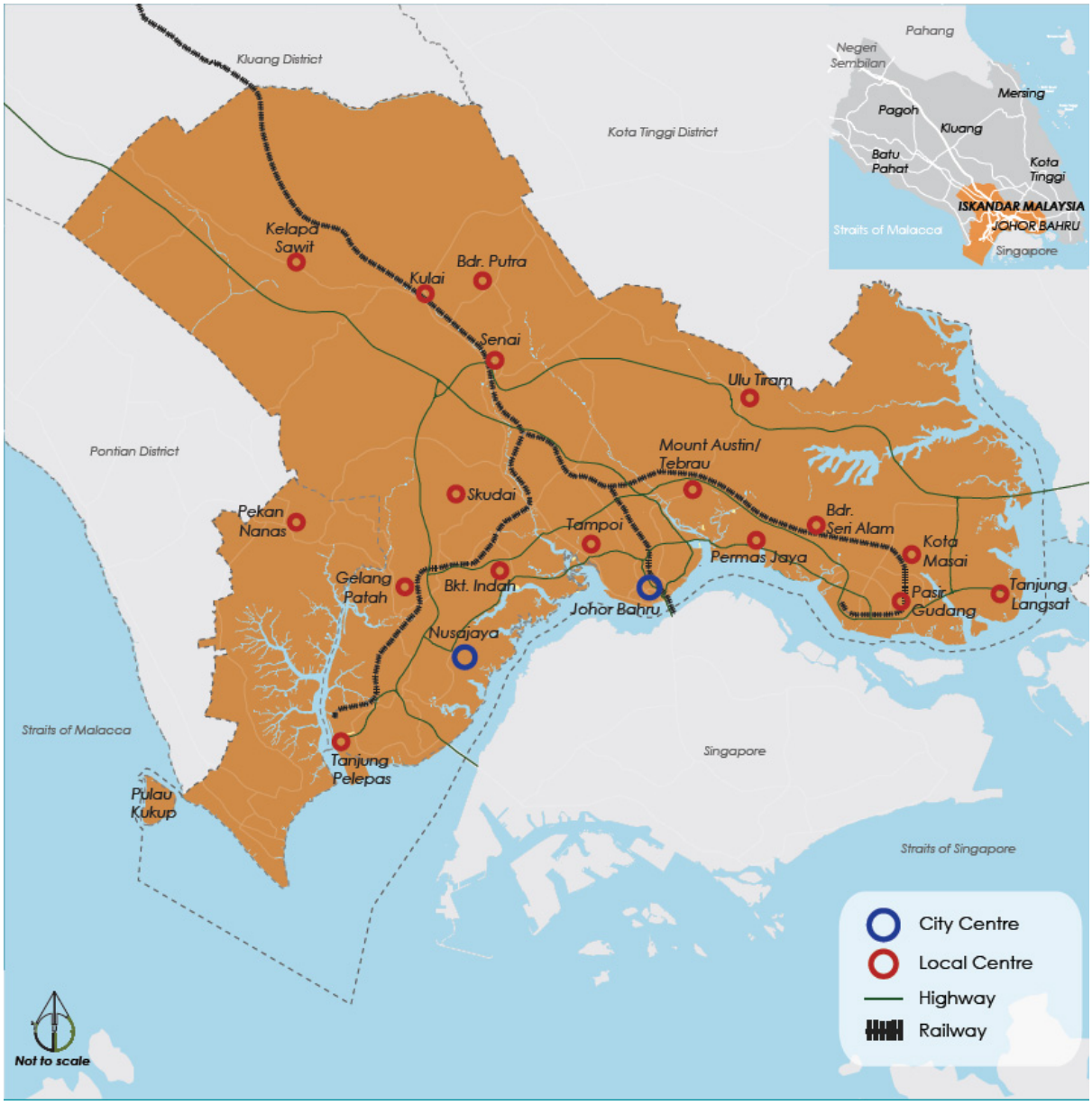

Figure 1 The Iskandar Malaysia project development region [3]

Data

To search for the potential locations to open a Chinese restaurant, the following data is required for the analysis:

- Number of existing restaurants (any) in the neighborhoods
- Number of Chinese restaurants in the neighborhoods, and their distances to the centres of the corresponding neighborhoods

The neighborhoods data is extracted from the geojson file published on the Github repository by leowmjw (https://github.com/TindakMalaysia/Johor-Maps) [4], as shown in the DataFrame in Figure 2. The database shows the administrative states and areas of the entire Johor state. The data used in the analysis is described in Table 1.

```
# Create a data frame of the neighborhood of entire Johor state

with open('Johor-DM-4326.geojson') as data:
    JB_geo_data = json.load(data)

Johor_par_list = []
df_Johor_par = pd.DataFrame
for i in range(len(JB_geo_data['features'])):
    Johor_par_list.append(JB_geo_data['features'][i]['properties'])

# Flatten the json file into a DataFrame
df_Johor_par = json_normalize(Johor_par_list)

# df_Johor_par.to_csv('Johor_Par.csv',index=False)
df_Johor_par.head()
```

	NAME	NAMA_DM	PAR_LAMA	DUN_LAMA	DM_LAMA	PENGUNDI	KODPAR
0	142/06/15	BANDAR CHAAH TENGAH	142	06	15	1123	142
1	155/34/10	BANDAR TENGGARA SELATAN	155	34	11	1608	155
2	165/55/11	BANDAR PEKAN NENAS SELATAN	165	55	13	3548	165
3	161/47/09	TAMAN DAHLIA	161	47	09	2743	161
4	157/38/04	PENAWAR	157	38	04	2006	157

Figure 2 The extracted data of the Johor state from the geojson file[4]

Table 1 Original attributes and their descriptions

No.	Attributes	Description
1	NAMA_DM	Name of the neighborhoods/administrative areas, (‘Neighborhood’)
2	PAR_BARU	Name of the belonging districts of the neighborhoods/administrative areas (‘District)
3	geometry.type	The type of the geometry border of neighborhood
4	geometry.coordinates	The coordinates plots of the border of neighborhood (‘Latitude’, ‘Longitude’)
5	PENGUNDI	Number of registered voters in the neighborhood, which are also local residents aged 21 or above (‘Number_of_Residents’)

The mean values of the *'geometry.coordinates'* are calculated and treated as the centre coordinates of the corresponding neighborhoods due to the unstable response from the geopy geocoder package. Additionally, to search for the nearby restaurants with respect to the centre of neighborhood, Foursquare API is employed in the analysis [5].

Reference

[1] D. o. S. Malaysia. Total Population by Ethnic Group, Local Authority Area and State, Malaysia [Online]. Available: https://web.archive.org/web/20131114184213/http://www.statistics.gov.my/portal/download_Population/files/population/04Jadual_PBT_negeri/PBT_Johor.pdf

[2] I. R. D. Authority. (2016). *Iskandar Malaysia*. Available: http://iskandarmalaysia.com.my/

[3] I. R. D. Authority. (2016). *Our Development Plan*. Available: http://iskandarmalaysia.com.my/our-development-plan/

[4] leowmjw. (2016, 10 February). *Completed migration of Johor Shapefile*. Available: https://github.com/TindakMalaysia/Johor-Maps/blob/master/2016/MAP/MIGRATED/result/13-Johor-New-DM-4326.geojson

[5] F. Inc. (10 February). *Foursquare Developer: Create Magical Real-World Moments for Your User*. Available: https://developer.foursquare.com/