A proper location for opening a Korean restaurant in city of Singapore

L. Zhou

Abstract

In this work, data science knowledge and skills are applied to solve a specific problem, i.e., to assist a business investor in deciding a proper location for opening a restaurant in Singapore. The decision is based on analyzing geographical data, business establishment and commercial activities in the areas, and many others such as political and social factors. Data collection and analysis are performed using various tools to address the problem. Foursquare services are applied to explore the surroundings of the interested areas, to find the relevant information of nearby restaurants and bars. Machine learning K-Means method is used to cluster the similar areas to identify an area with unique features.

1. Introduction

Singapore, city-state located at the southern tip of the Malay Peninsula, about 137 kilometers or 85 miles north of the Equator. It consists of the diamond-shaped Singapore Island and some 60 small islets; the main island occupies all but about 18 square miles of this combined area.

Singapore is a major financial and shipping hub, consistently ranked the most expensive city to live in since 2013. Singapore is also a popular tourist destination, with well-known landmarks such as the Merlion, Marina Bay Sands, Gardens by the Bay, the Jewel, the Orchard Road shopping belt, the resort island of Sentosa, and the Singapore Botanic Gardens, the only tropical garden in the world to be honored as a UNESCO World Heritage Site.[1]

The beautiful "Garden city" Singapore is full of commercial opportunities, however there are tough challenges for any business investors to open new business there.

This work is to leverage the Foursquare location data to explore or compare neighborhoods to assist a Korean business investor to choose a good location to open a Korean restaurant in Singapore.

This report is organized as follows. The specific problem to be solved in discussed in Section 2. Data collection and pre-processing is described in Section 3. In Section 4, methodology and data exploratory are presented. Analysis results are discussed in Section 5, and finally conclusion is drawn in Section 6.

[1] https://en.wikipedia.org/wiki/Singapore

2. Business Objective and Problem

The Korean investor is certain that Korean cuisine will be loved by the people in Singapore and plans to open a HIGH-END Korean restaurant there. The investor expects the restaurant to be located in a busy

business and commercial district with large people volume, including office workers and tourists, and without furious competition from nearby similar restaurants.

There are a number of administrative districts with established businesses, and many different types of restaurants catering various needs in those districts. To invest the restaurant the investor first needs to select the location by taking into account of commercial activities, business establishments, restaurant competition, and cost of operation in the area.

The problem of this work to be addressed is: deciding the proper location of the planned Korean restaurant that meets the investor's expectation.

3. Data collection and processing

To find a proper restaurant location, one needs to select a district among many possible choices, by collecting and analyzing the available information about the districts, i.e., the information on the surroundings, businesses and commercial activities, especially restaurants.

Singapore is divided into 28 districts under the old postal district system [2, 3]. Figure 1 shows Singapore Districts Map.

Figure 1. Singapore District Map

The districts are grouped into 3 regions: Core Central Region (CCR); Rest of Central Region (RCR); Outside Central Region (OCR). The three main regions of Singapore, constitute of several districts within them, which in turn contains the various towns or streets of the country. The grouping of Districts and towns in the Districts is shown in Table 1.

Table 1 lists the Group of Districts and towns in the Districts.

District	Code	Location
CCR	1	Boat Quay , Chinatown , Havelock Road, Marina Square, Raffles Place, Suntec City
CCR	2	Anson Road, Chinatown, Neil Road, Raffles Place, Shenton Way, Tanjong Pagar
CCR	6	City Hall, High Street, North Bridge Road
CCR	9	Cairnhill, Killiney, Leonie Hill, Orchard, Oxley
CCR	10	Balmoral, Bukit Timah, Grange Road, Holland, Orchard Boulevard, River Valley, Tanglin
CCR	11	Chancery, Bukit Timah, Dunearn Road, Newton
RCR	3	Alexandra Road, Tiong Bahru, Queenstown
RCR	4	Keppel, Mount Faber, Sentosa, Telok Blangah
RCR	5	Buona Vista, Dover, Pasir Panjang, West Coast
RCR	7	Beach Road, Bencoolen Road, Bugis, Rochor
RCR	8	Little India, Farrer Park, Serangoon Road
RCR	12	Balestier, Moulmein, Novena, Toa Payoh
RCR	13	Potong Pasir, Macpherson
RCR	14	Eunos, Geylang, Kembangan, Paya Lebar
RCR	15	Katong, Marine Parade, Siglap, Tanjong Rhu
RCR	20	Ang Mo Kio, Bishan, Braddell Road, Thomson
OCR	16	Bayshore, Bedok, Chai Chee
OCR	17	Changi, Loyang, Pasir Ris
OCR	18	Pasir Ris, Simei, Tampines
OCR	19	Hougang, Punggol, Sengkang
OCR	21	Clementi, Upper Bukit Timah, Hume Avenue
OCR	22	Boon Lay, Jurong, Tuas
OCR	23	Bukit Batok, Choa Chu Kang, Hillview Avenue, Upper Bukit Timah
OCR	24	Kranji, Lim Chu Kang, Sungei Gedong, Tengah
OCR	25	Admiralty, Woodlands
OCR	26	Tagore, Yio Chu Kang
OCR	27	Admiralty, Sembawang, Yishun
OCR	28	Seletar, Yio Chu Kang

CCR region is located in Singapore's main prime metropolitan area, of which owning any private residential property here would consider one as a 'high net-worth individual'. This is largely due to its locational attributes, its close proximity to the central business district (CBD), the quality of the properties, as well as the high property values as compared to the rest of the regions.

As the core central region has the busiest commercial activities and high-end residential properties, it has the most attractive reason to open a high-end restaurant in this region of 6 districts. Let's focus on and explore this core central region. Table 2 shows the CCR region with latitude and longitude data.

Table 2. CCR District latitudes and longitudes

District	Code	Latitude	Longitude
CCR	1	1.28895	103.84904
CCR	2	1.27929	103.85055
CCR	6	1.2925	103.85392
CCR	9	1.30374	103.83214
CCR	10	1.31355	103.79071
CCR	11	1.32363	103.81193

[2] https://www.singaporeexpats.com/housing-in-singapore/singapore-district-guide.htm \
[3] https://www.redbrick.sg/blog/singapore-districts-and-regions/

4. Methodology and Exploratory Data Analysis

Foursquare API and K-Means machine learning method will be applied in the data analysis. Foursquare API is applied to explore the place of interests, to acquire the business establishment information on the surroundings, especially the business related to eatery in the Districts in this project. Graphs are used to present the restaurant distribution in the Districts, and K-Means clustering method is applied to find the similarities among the Districts.

4.1 Venue exploratory using Foursquare API

The most common venues in CCR Districts within a radius of 500 meters is explored, shown in Table 3.

Table 3. The most common venues in CCR Districts

	CCR_district	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	1	Nightclub	Bar	Italian Restaurant	Hotel	Yoga Studio	Japanese Restaurant	Cocktail Bar	Café	Mexican Restaurant	Lounge
1	10	Chinese Restaurant	Bar	Bakery	Fruit & Vegetable Store	Spa	Seafood Restaurant	Food Court	Indian Restaurant	Dessert Shop	Department Store
2	11	Café	Shopping Mall	Asian Restaurant	Bakery	Indian Restaurant	French Restaurant	Malay Restaurant	Supermarket	Thai Restaurant	Noodle House
3	2	Café	Coffee Shop	Korean Restaurant	Cocktail Bar	Chinese Restaurant	Hotel	Food Court	Restaurant	Salad Place	Japanese Restaurant
4	6	Hotel	Café	Shopping Mall	Coffee Shop	Cocktail Bar	Event Space	Japanese Restaurant	Ramen Restaurant	Restaurant	Market
5	9	Boutique	Hotel	Sushi Restaurant	Café	Chinese Restaurant	Bakery	Shopping Mall	Bubble Tea Shop	Coffee Shop	Cosmetics Shop

The number of restaurants and bars in CCR Districts are shown in Figure 2.
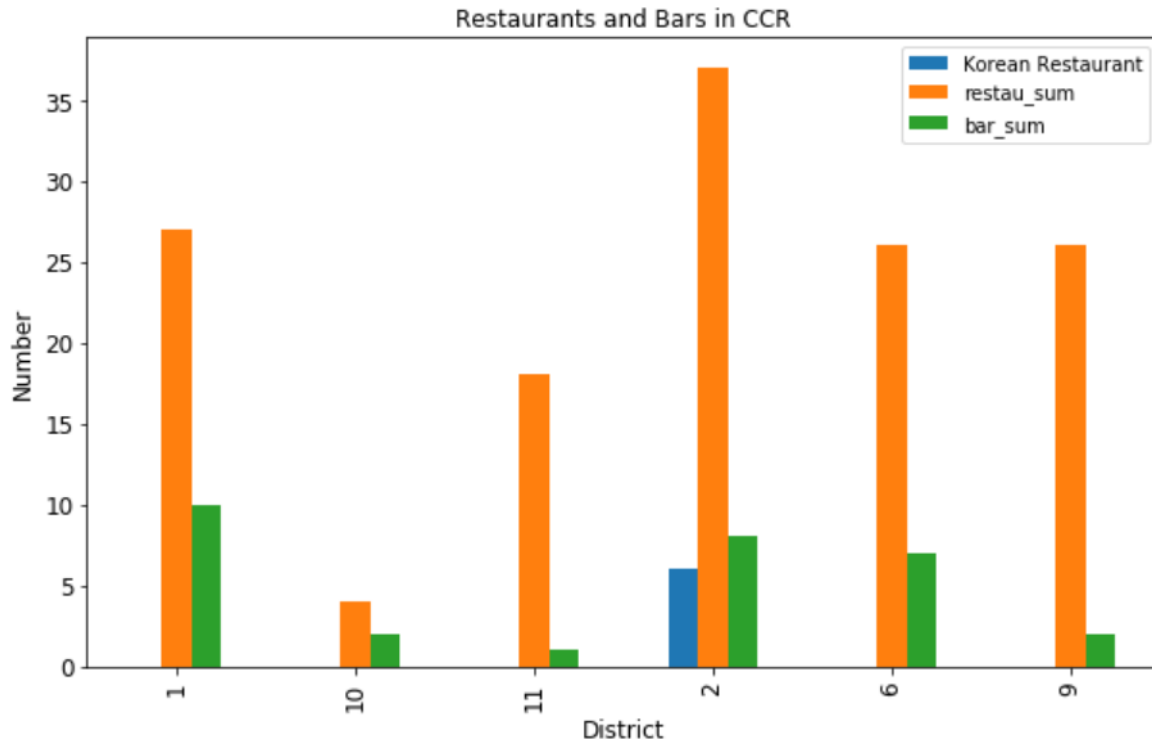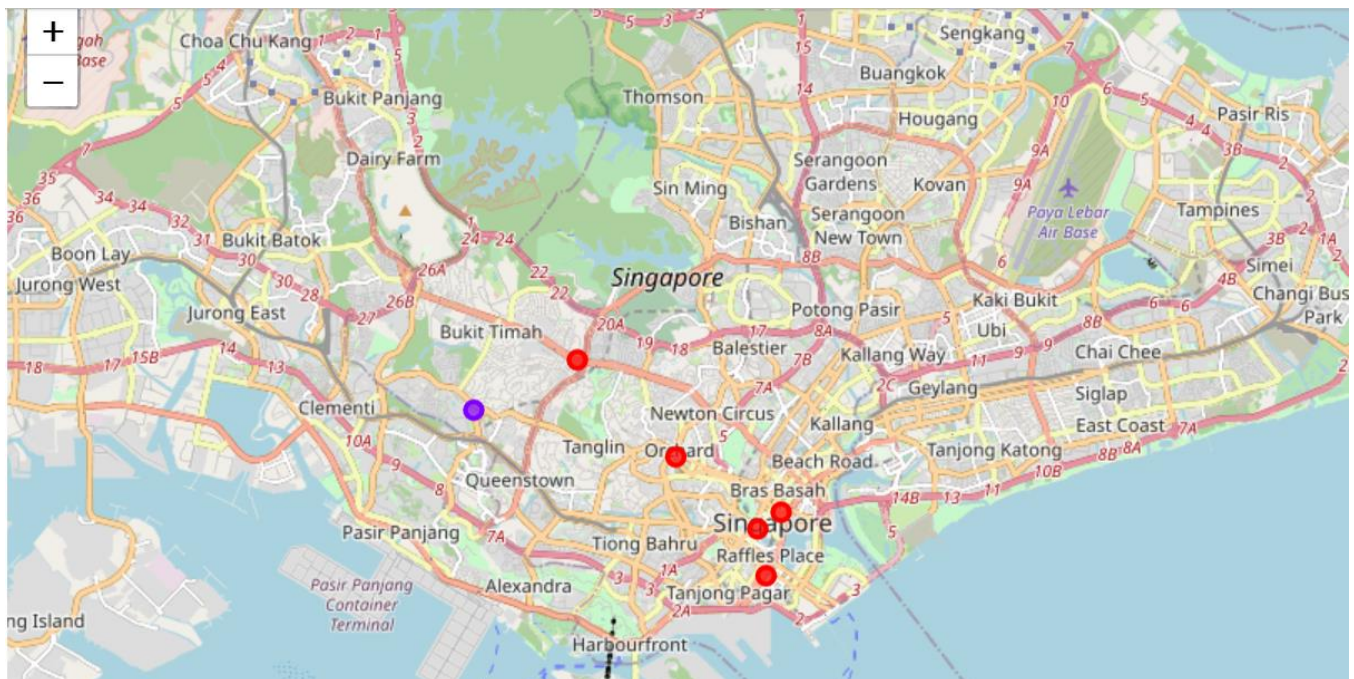
Figure 2. Restaurants and Bars distribution in CCR region

The 6 Districts in CCR region are grouped in 2 clusters via K-Means clustering method. The clustering is performed based on the information of most common venues in the Districts. Cluster 1: (10); and Cluster 2: (1, 2, 6, 9, 11).

5. Results and Discussion

Based on the analysis of the available information from public Websites and from Foursquare services, We recommend that the proper location of the proposed high-end Korean restaurant would be in CCR District 10.

First the CCR region has the most active commercial activities among the three regions in Singapore. A restaurant will naturally be part of it. Second, within CCR region there are many restaurants and bars, however, the number of restaurants in District 10 is the fewest (see restaurant distribution Figure), which means the competition may not be furious. Third, there are a few Bars operated here, which implies there are active social activities and therefore people traffic volume. Other Districts, with many restaurants and bars, have similarities and belong to the more established areas, especially District 11 area has high density private properties, so both the competition and operation cost will be high.

We recommend the Korean restaurant to be located in District 10. In the location selection process, Foursquare service provides a lot information on the surroundings, and helps in making the decision. However, the final decision should be made by analyzing more information, such as the scale of the established restaurants, the government blue plan, the ongoing business and commercial activities, distribution of population, and convenience of public transportation.

6. Conclusions

This work applied data science knowledge to solve a practical problem. It included problem statement and objectives of the work, and problem solving via data collection, pre-processing, and analyzation. The specific problem of this work is to assist a business investor in deciding a proper location for opening a restaurant in Singapore. A recommendation was made at the end of this work after going through the data collection and analysis work.

This work also helped to refine the learnt data science techniques and enhance our skill of applying data science knowledge to real problem solving.