

Distribution of Italian restaurants around Brisbane QLD Australia

Chee Too (Jason)

5/1/2020

1. Introduction

1.1 Background

Brisbane is the 3rd largest capital city of Australia and it is a culturally diverse city with 32.2% of residents were born overseas. According to Wikipedia, most of the Indian migrant concentrates at the north side of the city around Chermside area. Meanwhile, the southern suburbs and city area are home to a large proportion of Chinese, Vietnamese and Korean migrants

This project will look at the distribution of Italian restaurants around Brisbane area as there is no well-known suburb that has lots of Italian migrants or resident with Italian heritage.

When we consider all these problems, we can create a map and information chart where the we can visually inspect the distributions of the Italian restaurant.

1.2 Problem

This project would look for recommendation for someone who is looking to open Italian restaurant in Brisbane.

1.3 Interest

New entrant to the market would be interested to know the distribution of existing Italian restaurant in Brisbane. This could assist them in deciding the location of the new restaurant.

Local Italian restaurateur would be keen to know the distribution of Italian restaurants as well. Knowing how close/far the nearest competitor is important in deciding the business strategy, for instance: menu differentiation, price setting and customer segment targeting.

2. Data acquisition and cleaning

2.1 Data sources

Data are sourced for the following:

- i) Foursquare API
- ii) Zomato API

Data

2.2 Data cleaning

Data downloaded or scraped from Foursquare and Zomato are fit into data frame individually.

i) **Foursquare**

Foursquare API has a 50-results limitation applied. To bypass the restrictions, 2 centroids are used when requesting data – 1) Brisbane South - Mt Gravatt and 2) Brisbane North – Chermside.

Brisbane South and North are then merged into on data frame

ii) **Zomato**

Zomato API has a 20 results/query restriction apply. A loop is used to download 20 results each cycle until the maximum results reached, which is 80. Zomato API return empty result after the 80th.

There were a few irrelevant search results within foursquare result, like shoe store, embassy and language school. About 10 restaurants within foursquare result no longer trading but it was not reflected from foursquare. Furthermore, foursquare only return 37 results compared with 80 results from Zomato.

As the search result are more relevant and comprehensive from Zomato, the analysis will be based on data set scrapped from Zomato. There is no gain from merging the data set as foursquare results are basically a subset of Zomato's.

2.3 Feature selection

For each restaurant, the restaurant name, latitude and longitude are parsed from the dataset.

3. Clustering Modelling

Dataset is first modelled through K-means clustering using 3 clusters centroid.

K means clustering cluster the data based on its geographical features – north Brisbane, central Brisbane and south Brisbane.

This doesn't really help in analysing the distribution. Thus, I examine the data from another viewpoint: restaurant distance from CBD

4. Exploratory Data Analysis

Restaurants location are plotted using Folium map to give a visual of its distributions. Further exploratory data analysis is conducted. Restaurants are compared based on its distance from Brisbane CBD. It is found that the 50% of 80 Italian restaurants is within 3 km from Brisbane CBD and there is no other obvious location that have significant Italian restaurants.

This result coincides with the common knowledge of Brisbane as the dining precinct of Italian restaurants concentrate in the area of Brisbane City, South Bank and Fortitude Valley area.

5. Conclusions

It is significant that Italian restaurants are densely located in the city area, about 40 restaurants packed within 3 km range from CBD. Meanwhile, it is sparsely located in the outer suburbs at about 1-2 restaurants per suburbs.

In deciding the location of new Italian restaurant, business owner needs to consider its own competitive advantage and distinguished features compared with what already available. This is especially important when venturing into the City area (within 3 km from CBD) which is a highly competitive market. Otherwise, suburban location seems like a safer bet.

Discussions

As I mentioned before, there is no well-known suburb in Brisbane area that has significant Italian communities. Given the results of the data analysis, I reckon Brisbane residents just like Italian food in general and concentration around city area reflects customer volume.