Capstone Project - Finding the best area for opening Thai restaurant in Sydney

by

Thierrawat Jungprasert

April 5, 2020

1. Introduction

1.1 Background

The audience is the stakeholders in the Thai restaurant business looking to open a new Thai restaurant in Sydney so they would like us to help them find the best locations that match their criteria.

1.2 Problem

The audience problem is that Sydney is a big city and there are thousands of restaurants spread across the continent. It wouldn't be practical for the stakeholder to drive around each area and collect information on the restaurants. Using data science and mapping data from Foursquare, the world's leader in location service, combined with the Google API power to identify the coordinates of the location, would help to achieve this task in a sensible and most efficient way.

Therefore, we are conducting this business research in order to discover the best location for opening a Thai restaurant in Sydney.

The criteria for making evaluation are from the findings in our previous market survey which became the basis for our criteria for an evaluation in this analysis. The three findings are:

1. There's a strong relationship between the population density and the restaurant density in any particular area. The higher population density the higher density of the restaurant accordingly. This means that we can know the population density by knowing the restaurant density in the area.

2. In the survey about food preferences, it was concluded that Thai food is the most popular food in Sydney and people will choose Thai food over other food 95% of the time.

3. From the survey responses, people think that Malaysian food is very similar to Thai food that it can be a substitute for each other. 90% of the respondents say

that they are indifferent when it comes to choosing between Thai and Malaysian food.

Given the above findings, the criteria for finding the best area for opening a Thai restaurant in Sydney will be:

1. The area must not already have either Thai or Malaysian restaurant nearby.

2. Given the first condition is met, the area with more number of restaurants will be more promising than the area with a lower number of restaurants.

We will use python code to extract the information necessary for our analysis. We will then normalise and visualise the data to show all the areas candidate that is optimal for opening a Thai restaurant.

1.3 Interest

Using the result of the analysis, the stakeholder would be able to find the optimal location opening Thai restaurant. The cost of setting up business will be spent wisely on the location that will best earn them profit due to a more favorable environment of competition.

2. Data acquisition and cleaning

2.1 Data sources

To solve the business problem, we would be interested to know:

- number of existing restaurants in the neighborhood (any type of restaurant)
- number of existing Thai and Malaysian restaurants in the neighborhood, if any
- The distance of a certain location to Thai or Malaysian restaurants in the neighborhood, if any

we use Google API to obtain the latitute, longitute and addresses of the area we are evaluating. The data regarding the restaurant such as location, name and type will be from Foursquare.

2.2 Data cleaning and feature selection

There wasn't much effort on cleaning the data as the data from Google API was quite completed and in the format we need. The only formatting work we've done was cutting the country name 'Australia' from the full name of the address to avoid redundancies and make the data look cleaner. Here's the dataframe containing address of each area's center.

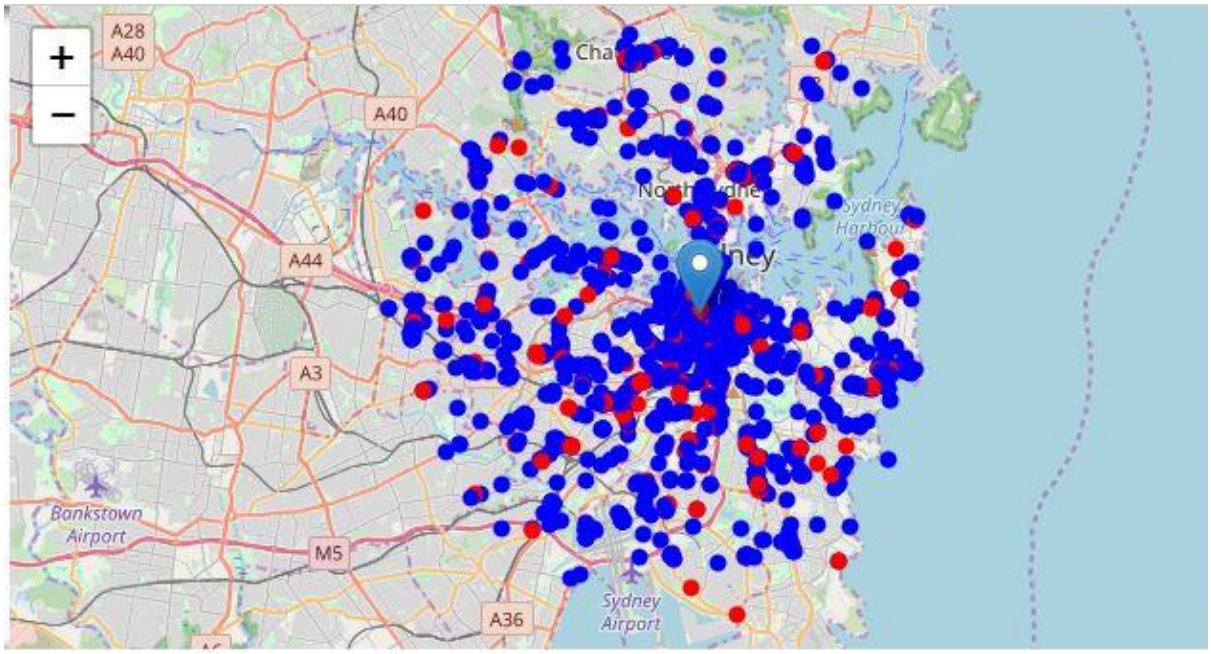
	Address	Latitude	Longitude	X	Y	Distance from center
0	1 Addison Rd, Manly NSW 2095	-33.811318	151.281864	4.673447e+06	-1.524991e+07	11984.990613
1	Reef Beach Track, Balgowlah Heights NSW 2093	-33.807150	151.272504	4.674647e+06	-1.524991e+07	11680.753400
2	24 Condamine St, Balgowlah Heights NSW 2093	-33.802983	151.263147	4.675847e+06	-1.524991e+07	11494.346436
3	5 Plant St, Balgowlah NSW 2093	-33.798815	151.253791	4.677047e+06	-1.524991e+07	11431.535330
4	65 Peacock St, Seaforth NSW 2092	-33.794647	151.244438	4.678247e+06	-1.524991e+07	11494.346436
5	79A Gurney Cres, Seaforth NSW 2092	-33.790479	151.235086	4.679447e+06	-1.524991e+07	11680.753400
6	7 Bampi Pl, Castle Cove NSW 2069	-33.786310	151.225736	4.680647e+06	-1.524991e+07	11984.990613
7	Hole in Wall Track, Manly NSW 2095	-33.824336	151.291582	4.671647e+06	-1.524887e+07	11711.532778
8	Hospital Rd, Manly NSW 2095	-33.820168	151.282220	4.672847e+06	-1.524887e+07	11208.925015
9	Balgowlah NSW 2093	-33.816000	151.272860	4.674047e+06	-1.524887e+07	10816.653826

3. Methodology

After retrieving the list of all restaurants in the area and their categories, we counted the number of them, both in terms of the total number of restaurant and only the restaurants that are either Thai or Malaysian restaurants. We found that Thai and Malaysian restaurants account for 13.93% of the total restaurants in Sydney.

Also, we found that the average number of restaurants in the area radius of 600m are 1.37. So, later, when it comes to selecting the minimum parameter of restaurant density for our criteria, we think that the density should be at least 5 restaurants in each area to satisfy the stakeholders' expectation as per the discussion with the stakeholders early on.

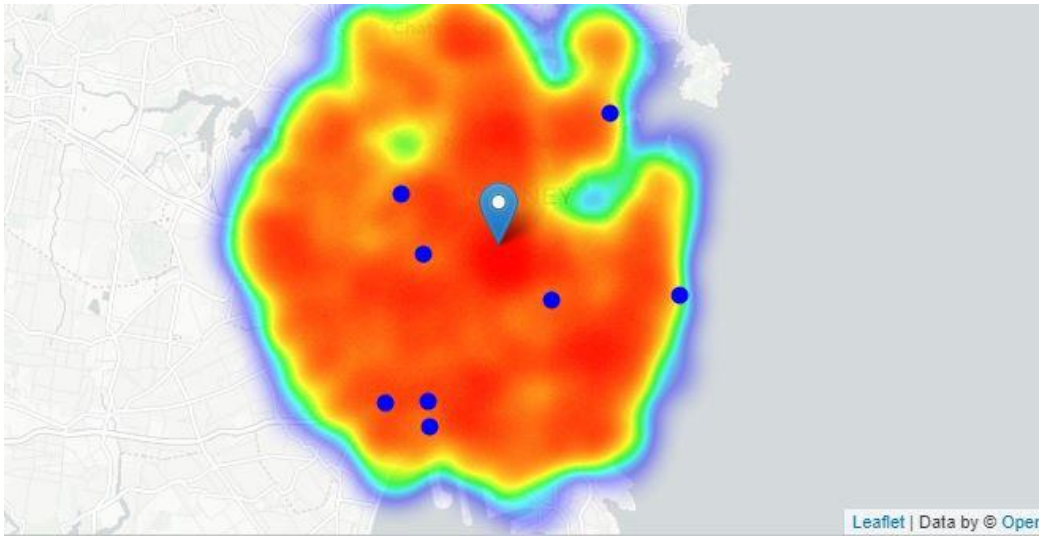
The total number of restaurants in the whole area is 1838, with 256 of them being Thai or Malaysian restaurants. We depicted these figures into the map with red dot representing the Thai or Malaysian restaurant and the blue dots representing the other kind of restaurants. Seeing the map below, it appears that Thai and Malaysian restaurant existed in most of the area in Sydney. This is not a surprise due to its popularity confirmed by our survey. Therefore, being able to the find the crowded area that does not have the Thai and Malaysian restaurant yet would be a great opportunity in doing a successful restaurant business.



Then, we calculate most important things for each location candidate: number of restaurants in a certain vicinity using radius of 600 meters and distance to closest Thai or Malaysian restaurants using the distance of 1 km. Here's what we found

- Locations with more than five restaurants nearby: 93
- Locations with no Thai and Malaysian restaurants within 1km: 163
- Locations with both conditions met: 8

Then we depict the above figure on a map. Now the blue dot represents the center o f each area that meet all the criteria mentioned above while the heat represents the total number of restaurants in the area. So we have a few locations having more than 5 restaurants in its radius of 600 and no Thai/Malysian restaurant in its radius of 1 k m. These locations are a potential candidate for a new Thai restaurant to be opened.



4. Results and Discussion

We have identified 8 potential areas that meet with our criteria for selecting location. However, which one of them will be the best one will depend on the other factor and the preferences of stakeholders. For example, if the stakeholders want to focus the high-income customer, then the wealthy area like Mosman would be the best option. On the other hand, if the focus is instead on the working-class customers aged around 25-40, the area in the suburb like Wolli Creek in the southwest be more preferable.

The blue dots of the 8 candidates were presented along with the heat map representing restaurant density in the area. Given the criteria and everything else equal, the blue dot situated in the more 'heated' area would be considered more preferable.

5. Conclusion

In this project, we aim to identify the best location to open Thai restaurant given our criteria, which is based on the survey and study conducted prior to this project. That is to say, the good location to open Thai restaurant should be one in the area that have high density of restaurant (representing high population density) but no Thai and Malaysian restaurant in the area yet.

By using the google API to pinpoint the coordinates of different area in Sydney and by using data from Foursquare to identify the restaurants and their type, we are able to build the dataframe and generate the map that depicts the features we need for our analysis, namely restaurant density by heatmap and existence of direct competitors which are Thai and Malaysian restaurant within a certain proximity by the blue dots.

After all, the findings in this analysis will have to be used along with other factors depending on final preferences of the stakeholders, i.e. whether the restaurant would be high-end or mid-end.