The Battle of Neighbourhoods

Open A New Restaurant in Auckland(NZ)

Jun Li

12/07/2021Introduction

1. Introduction:

1.1 Background

Set up a new business is complicated and full of challenges, especially for opening a new restaurant. Because there are many different types of cuisine, and the catering industry is very localized, almost all their business is from residences from nearby neighbourhoods. Once the restaurant started, and the people from local don't like the ideas of the new restaurant, there won't be easy to change the style, to gain customer acceptance back. Then the investor will possibly suffer a considerable loss.

1.2 Problem

To avoid the worst-case scenario, the investor carry out an initial investigation is essential. This investigation has to be able to answer these questions; "Where should I open a new Restaurant? What kind of restaurant should I do?". However, it could be very clueless for the business investors to evaluate, especially those who had never worked in the catering industry. Therefore, most of the investors only depend on ideas from the business broker or the information provided by the other people in the food serveries sector. In most cases, the investor will end up without having a complete picture on their own. Hence, the business investment has become a gamble.

This project aims to use Data Science methodology and its tools to answer those questions raised above for the new investor in Auckland (New Zealand). This study will work on the data on the existing restaurants in terms of popularity in each suburb. Then, using those findings, cross-examine against restaurant data in Auckland. Hopefully, the outcome could find out the best kind of cuisine, which is likely successful in each suburb within the great Auckland region.

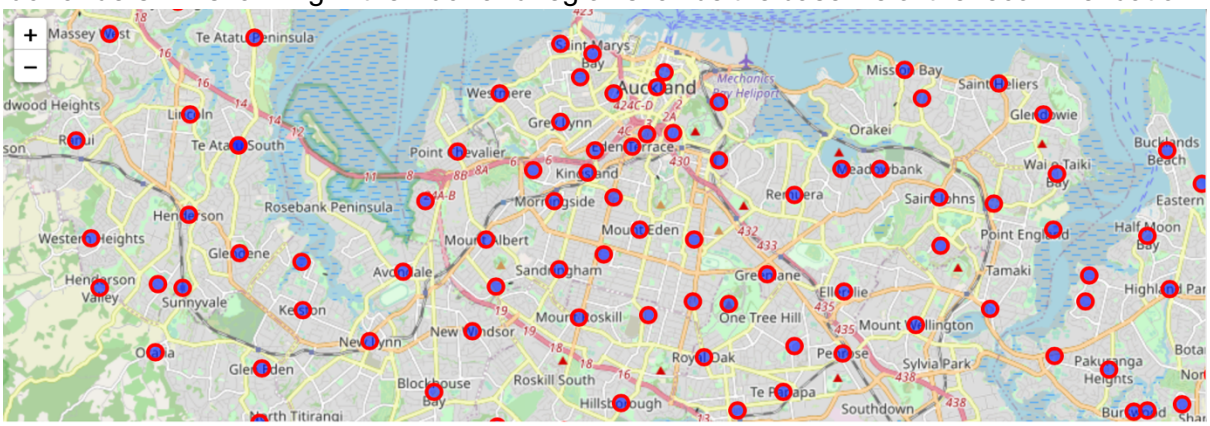
1.3 Interest

The output of this project will be a recommendation for the class of cuisine in the selected suburb. This finding could benefit a business investor who wants to start their first catering business.

1.4 The methodologic

The methodologic of this study is using the most updated numbers restaurant in one particular type of cuisine in the whole Auckland Region as a baseline to determine what kind of cuisine was lacking in the desired region. One cuisine's popularity is associated with the number of restaurants that have already been open in the area. Because, in a large area and long term, the supply and consumption will become ebullience by themselves.

Moreover, in the big Auckland region, the cuisine's popularity ranking was based on the data from 303 suburbs in 7 local boards. It should have been able to represent the general taste of Aucklanders. This ranking in the Auckland region shall be the baseline of the recommendation



study. The popularity in each suburb will be different to the overall ranking.

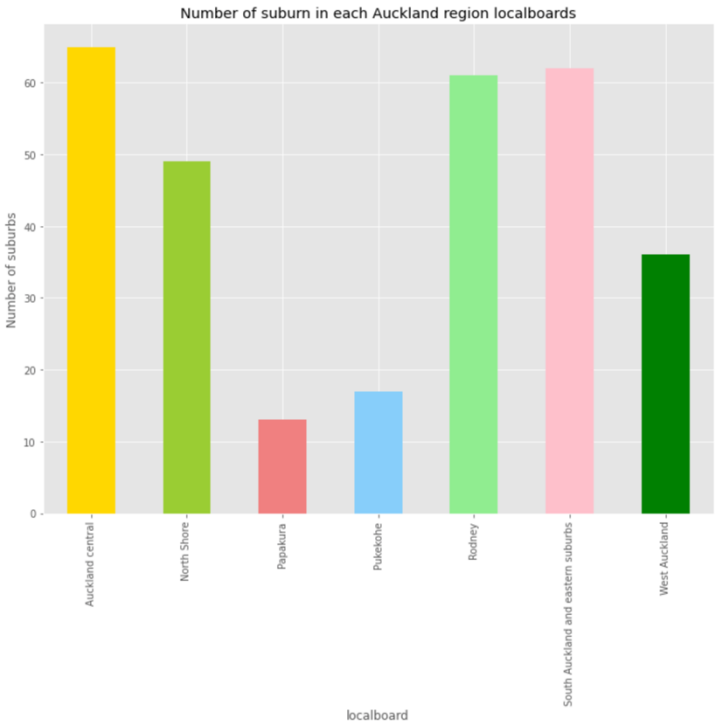
In the end, a recommended cuisine for a given suburb could be drawn by comparing a popularity cuisine ranking from a desired suburb against the baseline ranking. The degree shortage could be calculated through these analyses.

2 Data collection and data handling

2.1 Data sources

The fundamental data is the suburb in the big Auckland region. This data will be scrape from Wikipedia. List of Suburb Auckland:

"https://en.wikipedia.org/wiki/List_of_suburbs_of_Auckland



The suburb in each Auckland local board(Borough) then will be scraped into a dataset.

The following data required in the project is the coordinates of all suburbs in Auckland. The Geo-detector library in Python generated the coordinates for all suburbs. This library found the coordinates against the address given. Then the coordinates will be stored along with the suburb and local board belong.

Furthermore, all the coordinates of the suburb have been shown on the map of Auckland. They will also be treated as the central point of each suburb.

Foursquare is a server that could provide related geometric data, such as restaurants near the desired location. This data contains much information regarding geometric coordinates, cuisine, pictures and customer feedback, etc. This website provides an Application Programming Interface (API) for software and data scientists to utilize the service provided.

All the restaurants' information has been requested throughout the API against the central coordinates stored in a data frame. The restaurants are within one kilo-metre radius of the central coordinates. All the result is in JSON format.

2.2 Data cleaning

There is comprehensive information in the data returned from the Foursquare servers. The information includes the coordinates, name, category and feedback for venues within a range of the suburb centre. However, only name and venues category along with suburb and local board matters. Those information has been extracted from the JSON files.

There 3911 venues have been pickup from the serves. However, these venues included businesses from outside of the food and beverage sector. In this, those businesses will clean and deleted from the dataset

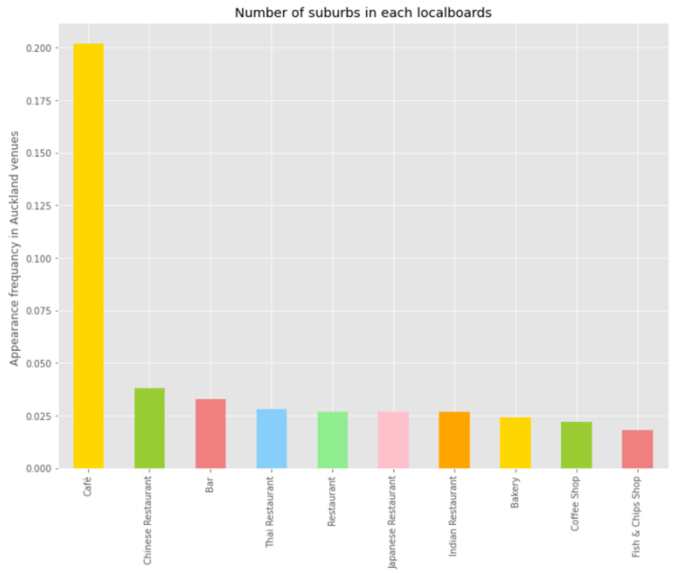
Nevertheless, some duplicate business venues have been returned from the request. This duplication error is due to the extensive search range around the centre point, and the search circle has overlapped. The duplication will be discovered and deleted during the data handling process.

There are 2986 venues left after the previous pre-processing.

3. Exploration and Data Analysis

3.1 Explore Big Auckland Region

In summary, the data set from the previous section have got 2986 venues, amount 303 suburbs. This dataset could be further cauterized according to the cuisine types. As a result, there are 248 different restaurants types, and they are in the order of amount that exists. However, not all the information is essential to the project. The scope of the project is to recommend the correct type of restaurant in a given area. As discussed in the methodology, only the top rank cuisine should be used as a baseline. The following table shows the top 10 restaurant classes in Auckland:



	venue	freq
0	Café	0.204
1	Chinese Restaurant	0.038
2	Bar	0.035
3	Japanese Restaurant	0.028
4	Bakery	0.026
5	Indian Restaurant	0.026
6	Restaurant	0.025
7	Thai Restaurant	0.025
8	Coffee Shop	0.024
9	Asian Restaurant	0.017

3.2 Explore Auckland Suburbs

Each restaurant has the suburb they belong to and the local board information in the venues data set. The local board information could categorize the restaurant into a group of local boards. Then, the cuisine type has been ranked according to the number that existed in each local board.

There is some exciting observation from comparing the rank in each local board to the Auckland region's overall ranking.

- The number one cuisine or catering business in number is Café.
- Almost all the top 5 catering types in each of the local boards also appear on the top 10 big Auckland cuisine ranking
- Not all top 10 cuisine class is shown on the Big Auckland cuisine ranking

Those are observation has established a good indication that the methodology of this project could produce a reasonable recommendation to an investor.

3.3 Explore Ponsonby

Ponsonby, Auckland, is an inner-city suburb just next to Auckland. It is one of the oldest suburbs with few fine dining restaurants. According to the traveller reviews report from

TripAdvisor, 'THE 10 BEST Auckland Neighbourhoods', Ponsonby sit in second place of the report. The recommendation will make base on Ponsonby current situation. It will be an ideal target sample for the project.

Foursquare API provided a packed venues list within a 1km radius from the centre of the Ponsonby neighbourhood. The figures showed 42 restaurants in 24 unique categories that have been picked up in the Ponsonby area.

4. Comparing outcome between Ponsonby and Auckland

```
----Ponsonby----
                   venue   freq
0                   Café  0.214
1     Italian Restaurant  0.071
2             Restaurant  0.071
3               Wine Bar  0.048
4          Shopping Mall  0.048
5           Burger Joint  0.048
6      French Restaurant  0.048
7    Japanese Restaurant  0.048
8            Pizza Place  0.048
9  Vietnamese Restaurant  0.024
```

```
 ----Auckland----
                   venue   freq
0                   Café  0.204
1   Chinese Restaurant  0.038
2                    Bar  0.035
3  Japanese Restaurant  0.028
4               Bakery  0.026
5    Indian Restaurant  0.026
6           Restaurant  0.025
7      Thai Restaurant  0.025
8          Coffee Shop  0.024
9     Asian Restaurant  0.017
```

4.1 Ponsonby and Auckland

By comparing the top 10 cuisines in Ponsonby with Overall Auckland figures, the result is very similar, especially in the leading five catering businesses in the list. At last, create a list for a selected suburb to see what catering has been missing comparing to the baseline.

- This process created a content-based recommendation system.
- The system will compare all existing catering classes already have with the top 10 in the Auckland region.
- Find which cuisine classes have been missing.
- Work out the degree of shortage.

The Auckland region restaurants sample is relatively large. Therefore the overall catering business categories ranking in Auckland can be used as the baseline.

At last, create a list for the selected suburb to see what catering has been missing comparing to the baseline. The list could the top catering classes have been missing with the degree of shortage. By comparing the top 10 cuisines in Ponsonby with Overall Auckland figures, the result is very similar, especially in the leading five catering businesses in the list.

4.2 Recommendation model

The recommendation is modulated by work out the difference between frequency radio $\{N_t\}$ for a particular cuisine class in a selected suburb and the baseline ratios $\{N_{baseline}\}$ for the same cuisine class, and different will divide by baseline ratios $\{N_{baseline}\}$. The equation is shown below.

$$\Delta C = \frac{(N_t - N_{baseline})}{N_{baseline}}$$

The result will be from a negative one to zero, or even more significant than zero. The negative means there is no such type of cuisine in the selected area at all. Zero means there are enough similar catering businesses that have been open in the area. If the result is larger than zero, there are over-open restaurants for this catering style, and the competition could be very high.

4.3 Recommendation list.

The result of the model is followed.

	rec_venue	dif
0	Chinese Restaurant	-1.000000
1	Indian Restaurant	-1.000000
2	Bakery	-1.000000
3	Coffee Shop	-1.000000
4	Bar	-0.314286
5	Thai Restaurant	-0.040000
6	Café	0.049020
7	Japanese Restaurant	0.714286
8	Pizza Place	1.823529
9	Restaurant	1.840000

5 Conclusion

Finally, the developed model has been tested on Ponsonby. From the list, there are four kind popular catering businesses was missing from the Ponsonby area. They are Chinese Restaurant, Indian Restaurant, Bakery, and coffee shop namely. Those restaurants should have less trouble to survive in Ponsonby since there is no competition in the area.

On the other hand, bar and Thai Restaurant will have moderate to high competition since the results showed there are very close to Zero.

Café, Japanese Restaurants, Pizza Places etc., are likely to have very high competition in the neighbourhood since the results from calculations are much higher than zero. There have a large number of similar restaurants existed. Therefore, the business investor who wants to start those types of catering should avoid Ponsonby.

6. Future action required.

There is simple and effective model has been developed from this project. However, there is more opportunity for improving this model. For example, the demography data can be embedded into the analysis. The social, economic and ethnic structure could have a significant contribution to the result.

The weight of each cuisine from different suburbs has not been taken into consideration. Even though the baseline data are from the whole Auckland region, it should present the general choice of restaurant for Aucklanders. However, the weight analysis could optimize the model and improve the accuracy of prediction.