

Where to open a new Vietnamese restaurant in Melbourne?: Neighbourhood Segmentation & Analysis

Tam Nguyen

Table of Contents

1. Introduction.....	1
2. Data Retrieval and Cleaning	2
2.1. Datasets.....	2
1. Geographical data	2
2. Melbourne venue data retrieved using Foursquare API:	2
2.2. Data Extraction and Cleaning	2
2.2.1. Geographical data	2
2.2.2. Venue data	3
3. Method.....	4
3.1. Specifying Candidate Neighbourhoods	4
3.2. Identifying Potential Neighbourhood	6
3.3. Clustering Potential Neighbourhoods.....	9
4. Results & Discussion	10
4.1. Results	10
Discussion.....	11
5. Conclusion	11

1. Introduction

Melbourne is an eventful and multicultural city - people from different cultures come together and build their homes here. The City of Melbourne is around 37 km², located in the central city area of Melbourne. Greater Melbourne's area of approximately 9900 km² of suburbs spread more than 40 km to the south, 30 km to the east, and 20km to the north.

In Australia, particularly Melbourne, the food service industry is thriving, making an impressive contributions to the national economy each year even though the majority of restaurants and café in the country are small businesses. While it is exciting to start a new restaurant business, there are a lot that need to be in the must-do list, such as design a menu, create a detailed business plan, invest in human assets, and the list can go on... Location is (*nearly*) everything to a business success. Even the best dishes and most resourceful staff cannot overcome a challenging venue that lacks potential regular customers or is overwhelmed by abundantly existing competitors, especially when the business is completely new. Many times it is very tempting to choose a busy and attractive spot in the area to set up the business, but even the busiest and most attractive of locations can hinder a successful business.

In this project, the **location to open a new Vietnamese restaurant in Melbourne** is the key aspect that I will strive to give recommendations on, using available datasets. Some criteria to identify potential areas are:

- not overcrowded with restaurants (no more than 1 restaurant within 300m)
- no Chinese/Vietnamese restaurants in close quarter (600m)
- close to the Melbourne city centre

2. Data Retrieval and Cleaning

2.1. Datasets

Based on the criteria of potential areas, some data types are to be retrieved:

- not crowded with restaurants → number of existing restaurants in a neighbourhood
- no Asian/Chinese/Vietnamese restaurants in close quarter (600m) → number of and distance to Vietnamese restaurants in the neighbourhood
- close to the central city → distance of a potential neighbourhood from city center

The data to be utilised in the project are:

1. Geographical data

`geopy` is used to retrieve, extract information about a location. `geopy` generates the coordinates of an address, and help obtain an approximate address given the location.

2. Melbourne venue data retrieved using Foursquare API:

Using extracted Melbourne postcodes and Foursquare API, data consisting of venues in Melbourne is retrieved. The retrieved data will be put into a dataframe, consisting of 7 columns:

Field	Description	Example
neighbourhood	name of the neighbourhood (suburb)	Melbourne City
neighbourhood latitude	latitude of the neighbourhood	-37.817403
neighbourhood longitude	longitude of the neighbourhood	144.956776
Venue	name of the venue	Royal Stacks
Venue latitude	latitude of the venue	-37.817867
Venue longitude	longitude of the venue	144.958489
Venue Category	category of the venue	Burger Joint

2.2. Data Extraction and Cleaning

2.2.1. Geographical data

Greater Melbourne area covers nearly 10,000 km²; while the city of Melbourne covers ~37km². For such a vast area, I need to narrow down the area to explore. In this project, the **initial candidate areas** are specified as neighbourhoods of 600m-radius which are within 12kilometres from Melbourne city centre (Melbourne CBD). Information about the initial neighbourhoods are stored in `df_locations` dataframe:

Field	Description	Example
Address	address of the neighbourhood	Jellicoe Street, Ivanhoe, City of Banyule, Vic..
Latitude	latitude of the neighbourhood	-37.817403

Longitude	longitude of the neighbourhood	144.956776
X	center X in Cartesian coordinates of the neighbourhood	4.976681e+6
Y	center Y in Cartesian coordinates of the neighbourhood	- 1.441943e+7
Distance from center	Distance from Melbourne CBD	11954.915307

2.2.2. Venue data

Given the **initial candidate areas**, Foursquare API was utilised to retrieve only venues which are categorised as “restaurant”, with keywords (restaurant, diner, steakhouse, taverna) in their venue category. Importantly, Asian/Chinese/Vietnamese-style restaurants in the initial area are also identified.

3. Method

This project attempts to identify areas of Melbourne that have low restaurant density, particularly those with low number of cafes. We will limit our analysis to area ~12km around Melbourne CBD. There are 3 main steps:

1. **Specifying candidate neighbourhoods.** Here, I collect data about *location and type (category) of every restaurant within 12km from Melbourne city center* (Melbourne CBD). Also, *Asian/Chinese/Vietnamese-style restaurants* are to be identified (according to Foursquare categorization).
2. **Identifying potential neighbourhoods.** Here, I calculate and explore restaurant density in the specified candidate areas. Heatmaps and clusters of restaurants will be used to (visually) identify some promising areas close to Melbourne CBD with *low number of restaurants in general (and no Asia-style restaurants in neighbourhood area)* and focus the attention on those areas.
3. **Clustering potential neighbourhoods.** Focusing on most promising areas, recommendation criteria are specified: locations with *no more than one restaurants in radius of 300 meters*, and locations *without Chinese/Vietnamese-style restaurants in radius of 600 meters*. We will present map of all such locations but also create clusters (using *k-means clustering*) of those locations to identify general areas / neighbourhoods / addresses which should be a starting point for final physical/street-level exploration and search for an optimal location to open a new Vietnamese restaurant.

3.1. Specifying Candidate Neighbourhoods

Firstly, a grid of candidate areas are created. The areas are equally spaced, centered around city center and within **~12km** from Melbourne city centre since Greater Melbourne is quite vast. The neighbourhoods will be defined as circular areas with a radius of **600** meters, so our neighbourhood centres will be **1.2 km** apart. The functions used to convert from latitude, longitude into Cartesian coordinates and vice versa are borrowed from the sample notebook. In this step, **314 candidate areas** are specified (as in Figure 1).

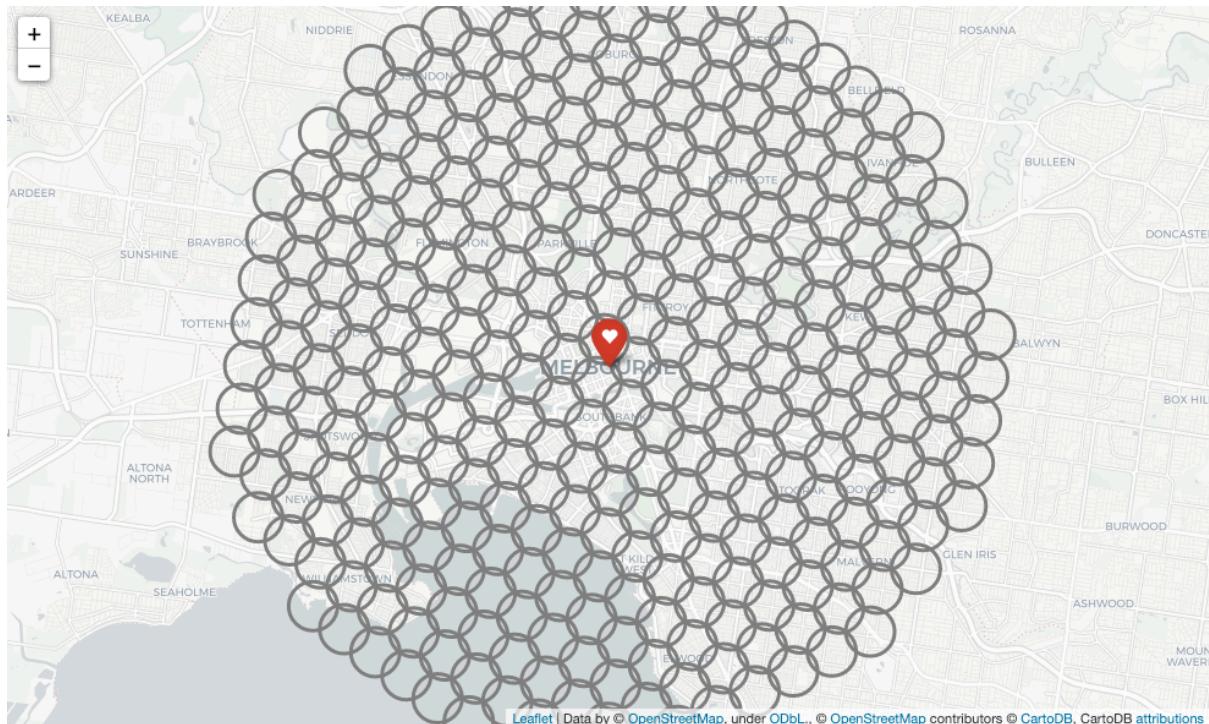


Figure 1. 314 candidate areas

Secondly, Foursquare API is utilised to retrieve restaurants in the 314 candidate areas. Asian/Chinese/Vietnamese-styles restaurants are also identified. The category IDs are taken from Foursquare website as follows:

- Food category: 4d4b7105d754a06374d81259
- Asian restaurant: 4bf58dd8d48988d142941735
- Chinese restaurant: 4bf58dd8d48988d145941735
- Vietnamese: 4bf58dd8d48988d14a941735

Here, the venues are found using Foursquare API with radius = 650 meters from each area centre so that no restaurant is missed. Information about the restaurants are then locally stored in `.json` files:

- `restaurants.json`: all restaurants found within the 314 areas
- `direct_competitors.json`: all Asian/Chinese/Vietnamese-style restaurants found within the 314 areas
- `location_restaurants.json`: restaurants in each area (radius = 600 meters)

Thirdly, exploratory analysis is performed including: calculation of the number of restaurants in every area candidate, calculation of the distance to nearest Asian-style restaurant from every candidate centre.

3.2. Identifying Potential Neighbourhood

Firstly, heatmaps & restaurant clusters are shown on maps to visually identify potential zones in the 314 candidate areas. Below are maps of: heatmap of all restaurants (Figure 2), heatmap of all Asian/Chinese/Vietnamese restaurants (Figure 3), cluster of Asian restaurants (Figure 4) in the candidate areas.

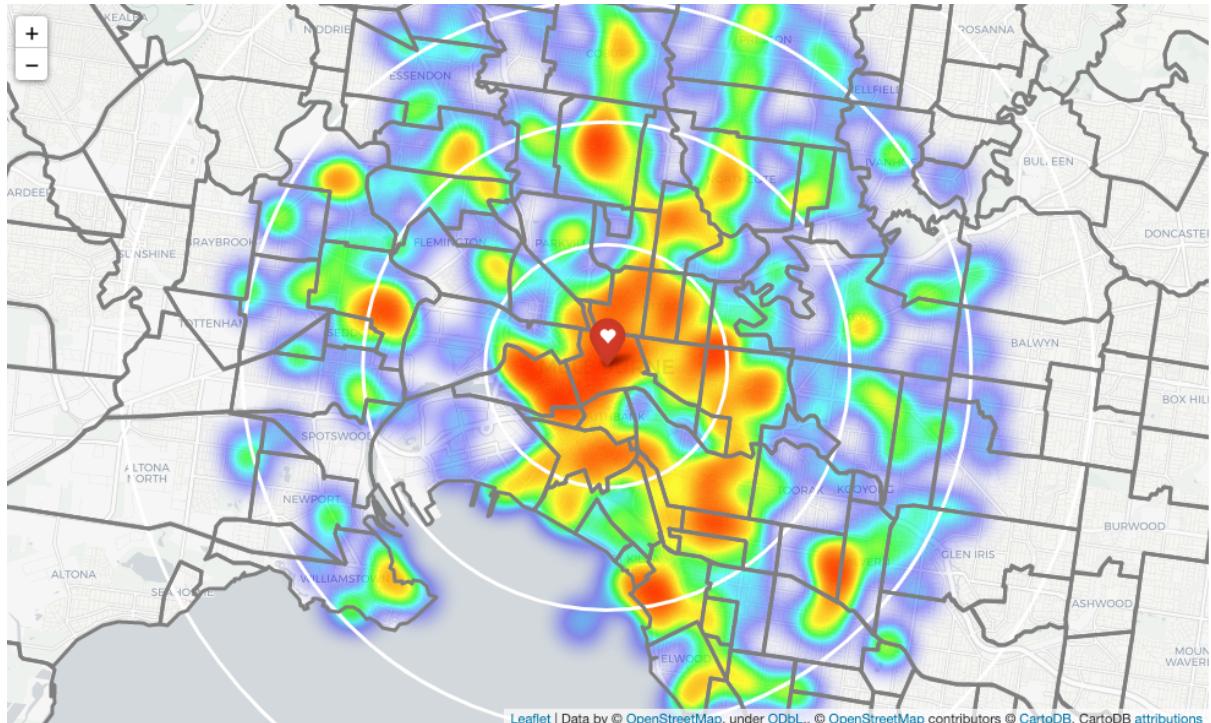


Figure 2. Heatmap of all restaurants in the candidate areas. 4 white circles indicate distance of 3km, 6km, 9km and 12km from Melbourne centre

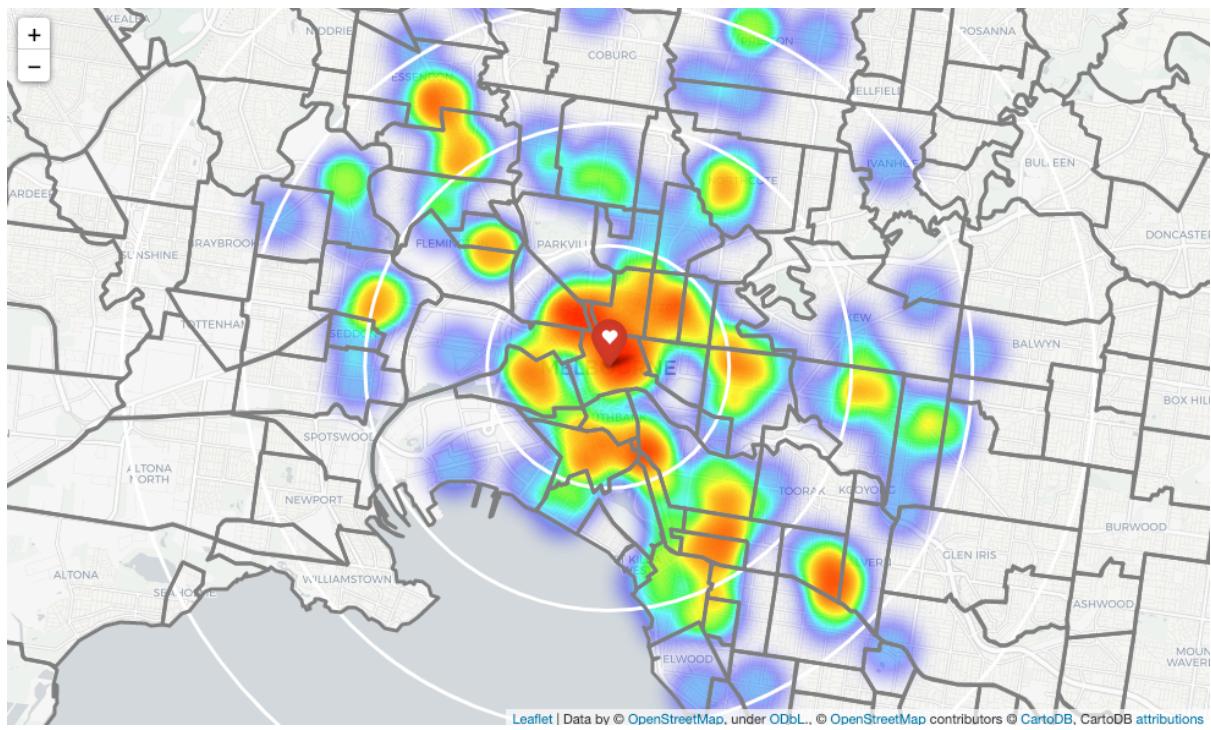


Figure 3. Heatmap of all Asian/Chinese/Vietnamese restaurants in the candidate areas

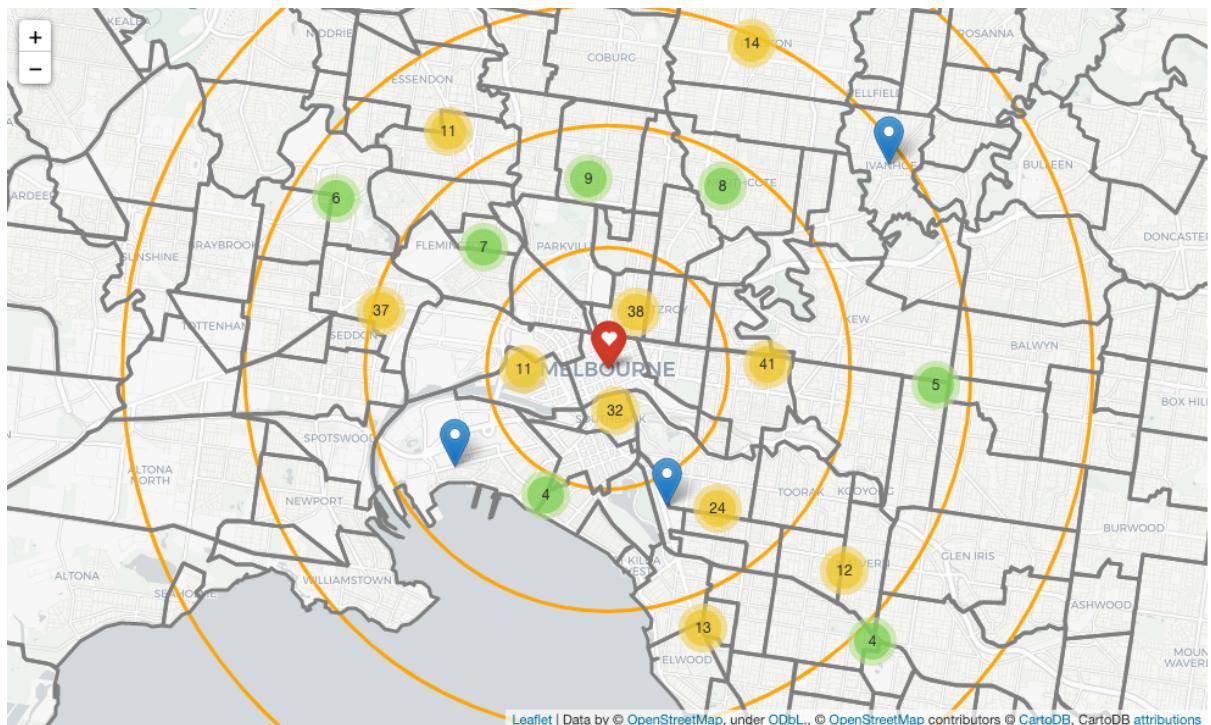


Figure 4. Clusters of Asian/Chinese/Vietnamese restaurants in the candidate areas

Initial observation: Areas with closest pockets of low Asian-style restaurant density positioned east, north, and north-east ~ 3km from city center. To be specific, the potential neighbourhoods are: **East Melbourne, North Melbourne, Clifton Hill (North-east), Parkville.**

Secondly, based on the initial observation above, we will now focus our analysis on areas *east, north, and north-east from Melbourne CBD* - we will move the center of our **area of interest** and reduce its size to have a radius of **4km** (as in Figure 5).

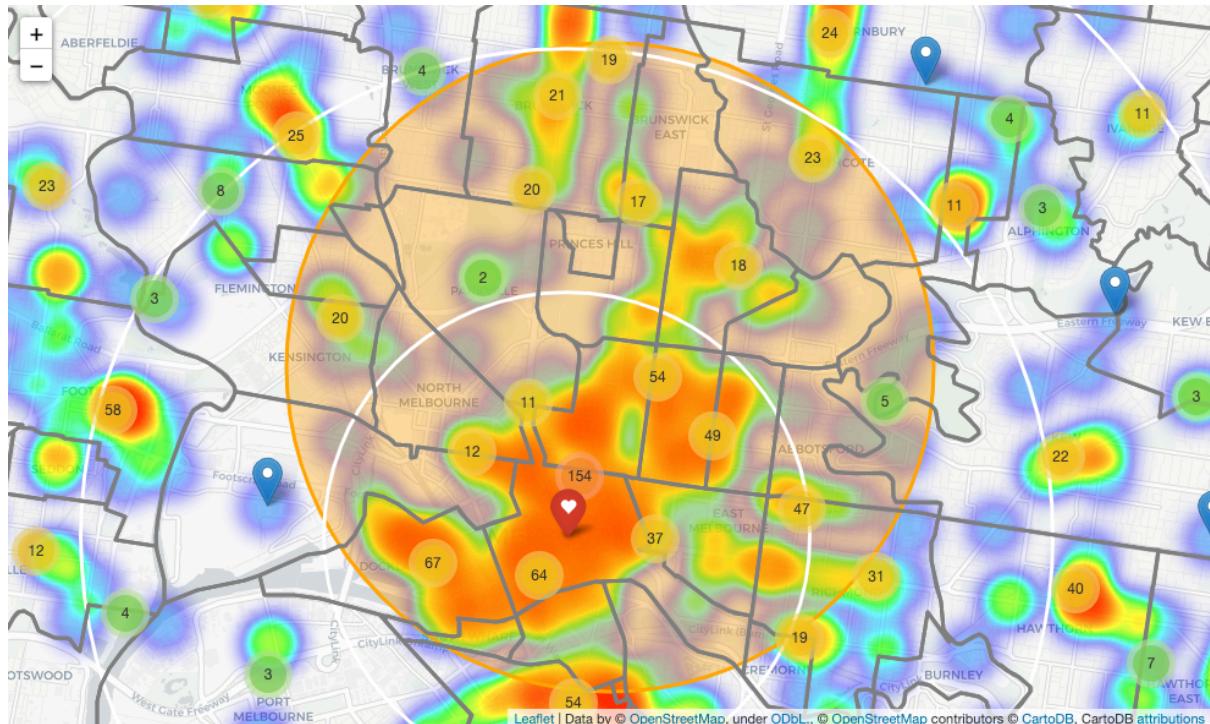


Figure 5. The narrow area of interest (north-east of Melbourne CBD, radius = 4km)

Thirdly, smaller locations within the **area of interests** are specified in order to narrow down the potential locations. The locations are 300m apart. 642 locations are created. For each location, the **number of restaurants in radius of 300m** and **distance to closest Asian restaurant** are calculated.

Fourthly, these 642 locations are filtered based on 2 criteria:

- Locations with no more than 1 restaurant in radius of 300m
- Locations with no Asian restaurant in radius of 600m

280 locations satisfy the conditions. They form the highly potential areas to set up a new Vietnamese restaurant. Below is the locations on the map:

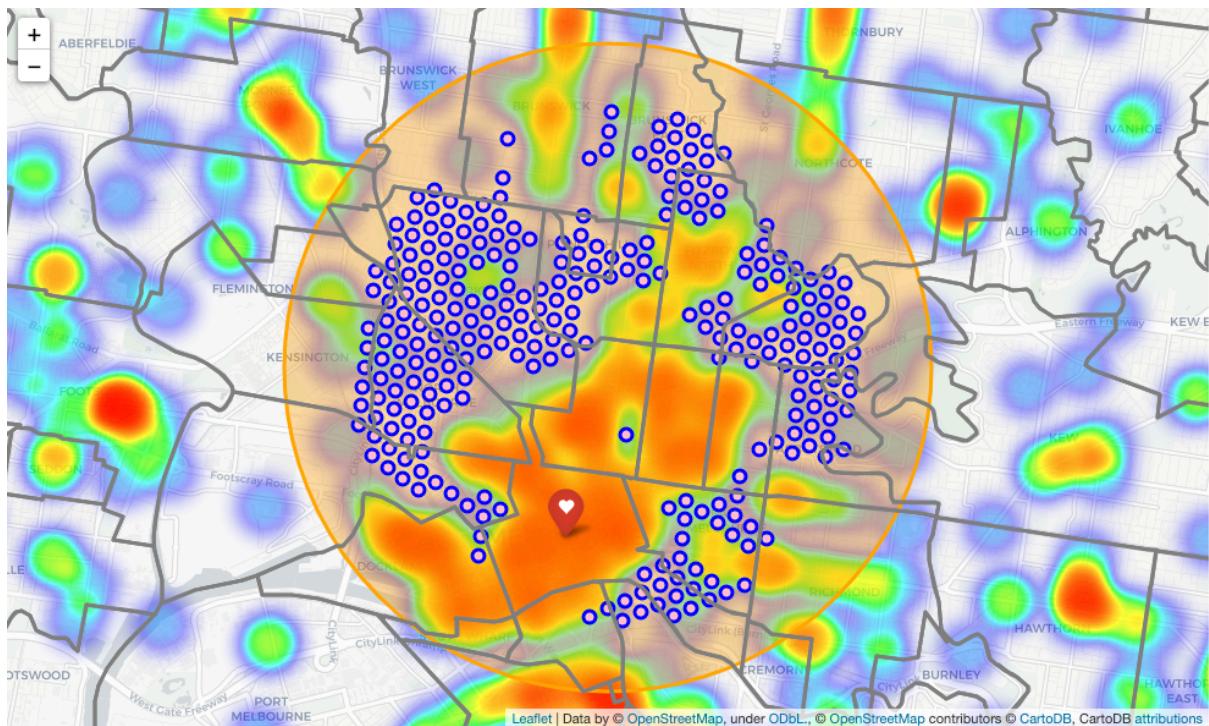


Figure 6. 280 locations satisfying 2 filtering conditions on map

In the next step, those locations are clustered to create centres of zones containing good locations.

3.3. Clustering Potential Neighbourhoods

The 280 locations are clustered using k-Means to create 5 groups. Figure 7 shows the 5 clusters on map:

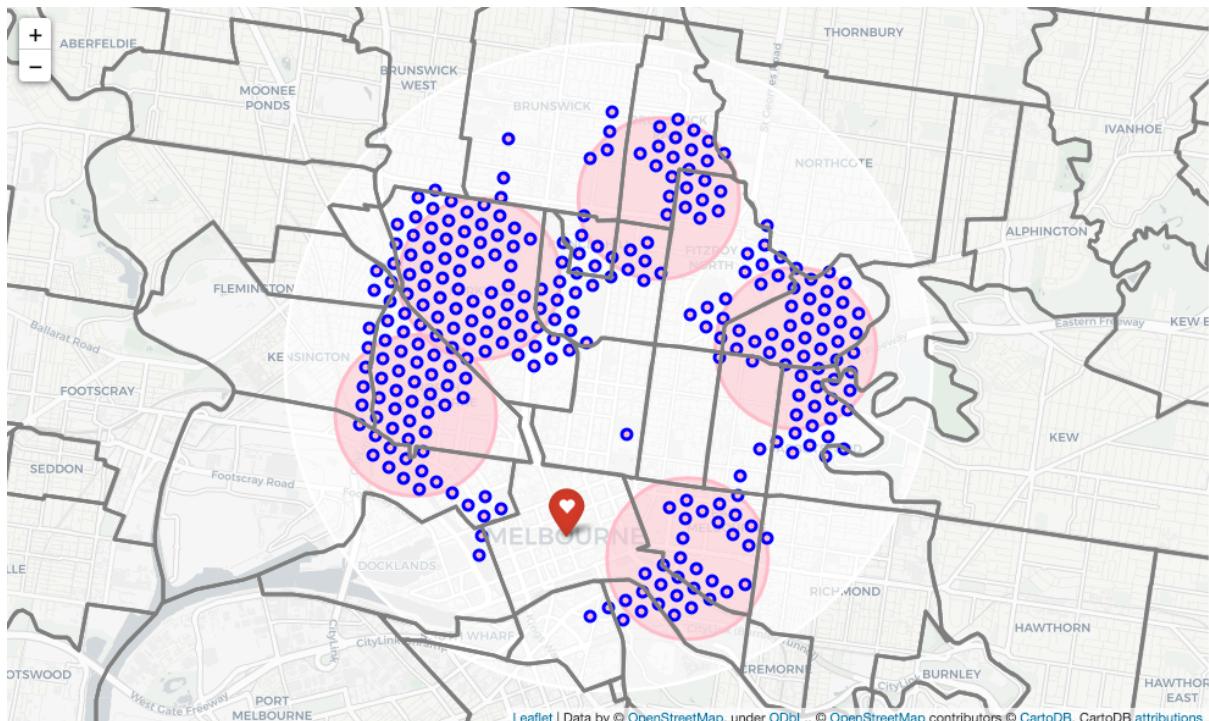


Figure 7. 5 clusters on map

The 5 clusters represent groupings of most of the candidate locations and cluster centres are placed in the middle of the zones with location candidates. 5 clusters identified:

- North Melbourne (-37.7928, 144.9413)
- Parkville (-37.78333, 144.95)
- Brunswick East x Fitzroy North
- Clifton Hill (-37.79798, 144.99533)
- East Melbourne (-37.8166700, 144.9879000)

4. Results & Discussion

4.1. Results

From the 5 clusters, the coordinates and address of the cluster centres are obtain. They are the ones which can be presented to stakeholders. Below is the 5 locations on map:

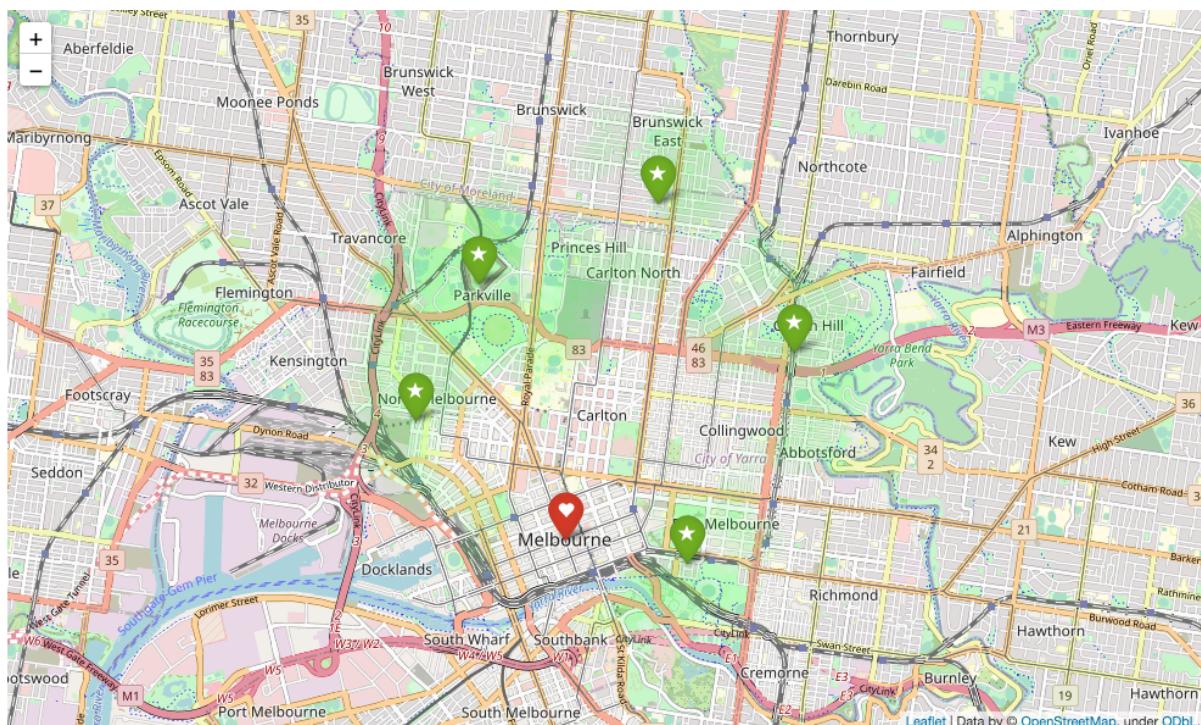


Figure 8. 5 recommended locations for initial exploration

The distances from these locations to Melbourne city centre are also calculated:

Address	Distance to Melbourne CBD
12, John Street, East Brunswick Village, Brunswick East, City of Moreland	5.4km
Melbourne Zoo, Poplar Road, Parkville, City of Melbourne	4.2km
Parslow Street, Clifton Hill, City of Yarra	4.6km
Wellington Parade South, East Melbourne, City of Melbourne	1.9km

Discussion

In part 3.1, 314 areas which are within 12km from Melbourne city centre are specified. These areas all have a radius of 600 meters.

In part 3.2, all restaurants, including Asian/Chinese/Vietnamese restaurants within the 314 areas are retrieved using Foursquare API. Total number of restaurants: 1549 Total number of potential Vietnamese restaurant competitors: 281 Percentage of competitor restaurants: 18.14%.

Using heatmap & Asian restaurant cluster on map, potential areas are identified, including: East Melbourne, North Melbourne, Clifton Hill (North-east), Parkville. From that, the size area of interest was narrowed down to 4km radius, and its centre was moved to the north of Melbourne centre.

Within the 4km-radius area, 642 candidate neighbourhoods (300m radius) are specified. The criteria of good neighbourhoods are: no more than 1 restaurant in 300m radius & no Asian restaurant in radius of 600m. Therefore, the number of restaurants in 300m radius and distance to closest Asian restaurant was calculated for each candidate. 280 locations were obtained which meet the 2 criteria.

In part 3.3, k-Means was applied on the 280 locations, generating 5 clusters. The clusters closely match the potential areas identified in part 5. Distance from the centre location of the 5 clusters to Melbourne city centre was calculated. The results show: Wellington Parade South, East Melbourne is the closest (1.9km from Melbourne CBD); while 12, John Street, East Brunswick Village is the furthest (5.4km).

5. Conclusion

The results can be used by the stakeholder as preliminary recommendations for further physical exploration of the potential neighbourhoods.

Limitation & Further research:

Venue information retrieved using Foursquare, so results generated is dependent on the data available, which can be not so up-to-date.

More criteria can be included, such as: characteristics of the neighbourhoods, surrounding environment, foot traffic, parking meters, rental prices, etc.