

COURSERA CAPSTONE REPORT FOOD MULTICULTURISM IN MELBOURNE

SWATHI RAYAPU

Table of Contents

Introduction.....	2
Background.....	2
Problem statement.....	2
Stakeholders.....	2
Data.....	2
Data Acquisition.....	2
Data Selection.....	3
Methodology Section.....	4
Foursquare Places API.....	4
Exploratory Data Analysis.....	4
Machine Learning.....	6
Results Section.....	8
General Observations.....	10
Recommendations for Businesses.....	10
Conclusion Section.....	10
References.....	10

Introduction

Background

Melbourne is a multi-cultural city with almost 140 cultures represented. People from countries like China, Vietnam, Italy and Greek migrated in early years and contributed significantly to the city's culture. Whereas, recent migrants are the large numbers of international students and IT industry employees from countries like India, Thailand, Malaysia and African countries.

This diversity reflects in many aspects of everyday life and activities. According to City of Melbourne resident and ethnicity profiles, "56% of people who live in inner city suburbs are overseas born and 48% of people speak a language other than English at home". These statistics show that Inner Melbourne is a multi-cultural city. To validate this statement further food scene can also be analysed.

Problem statement

To examine Melbourne's food choices, distribution of restaurants with cuisines of different countries. To explore which suburbs are similar in terms of food offerings by segmenting them into clusters. To validate whether availability of cuisines coincide with migrant patterns. Finally decide if Melbourne is truly a multicultural food city.

Stakeholders

This analysis can be used for business purposes and general understanding as well. Stakeholders for this analysis:

1. Companies who are in food business/individual or group of people who want to open a restaurant with a particular cuisine: They can learn about various food related factors about an area like for which cuisine the area is famous for, what other cuisines are available in the area, what are other areas similar to this area etc. This enables them to know if an area is over saturated with a type of cuisine or if there is still scope for new restaurants and choose an area.
2. Migrants/General public: This analysis provides the overall food scene of the city. They'll understand which suburb has which kind of restaurants. It can help them to make few decisions related to where they want to live or simply where they should go if they want to taste a cuisine.

Data

Data Acquisition

- Initial Dataset: An open source dataset which had Australian cities data has been used to filter for Inner Melbourne data. Inner Melbourne has 3 cities, City of Melbourne, City of Port Phillip and City of Yarra. These 3 cities are extracted with their suburbs and postal codes. After filtering, the initial dataset to start the analysis has 3 columns. The data frame has 3 cities and 35 suburbs.

(35, 3)

	Postcode	City	Suburb
0	3000	Melbourne	CBD
1	3002	Melbourne	East Melbourne

- Location Coordinates extraction: To get location coordinates for the above-mentioned file and suburbs, Google Geocoding API has been used. After processing for coordinates, postcode is no more required so removed it. Data set has 4 columns as mentioned below:

(35, 4)

	City	Suburb	Latitude	Longitude
0	Melbourne	CBD	-37.81110	144.97021
1	Melbourne	East Melbourne	-37.81132	144.97799

- Venues extraction: To get restaurant venues for each of the coordinates in the above file, Foursquare Places API has been used. After extracting venues, the final dataset has 7 columns as mentioned below:

(1720, 7)

	Suburb	Suburb Latitude	Suburb Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	CBD	-37.812245	144.962164	Xing Fu Tang	-37.812206	144.962066	Bubble Tea Shop
1	CBD	-37.812245	144.962164	Brother Baba Budan	-37.813445	144.962137	Coffee Shop
2	CBD	-37.812245	144.962164	Emporium Café Court	-37.811925	144.963785	Food Court
3	CBD	-37.812245	144.962164	Calia	-37.812724	144.963930	Japanese Restaurant

Data Selection

As the area of interest is to see the availability and distribution of various cuisines across suburbs, few venues that are irrelevant to the analysis are removed from the final dataset. For example:

- Common venues as Ice Cream Shop, Café, Coffee Shop, Juice Bar
- Pub, Bar, Brewery

After all the data preparation, the final dataset has 55 unique food categories and 349 unique values.

```
: print('There are {} unique FOOD categories.'
```

```
There are 55 unique FOOD categories.
```

```
: print('There are {} uniques venues.'.format
```

```
There are 349 uniques venues.
```

Methodology Section

Foursquare Places API

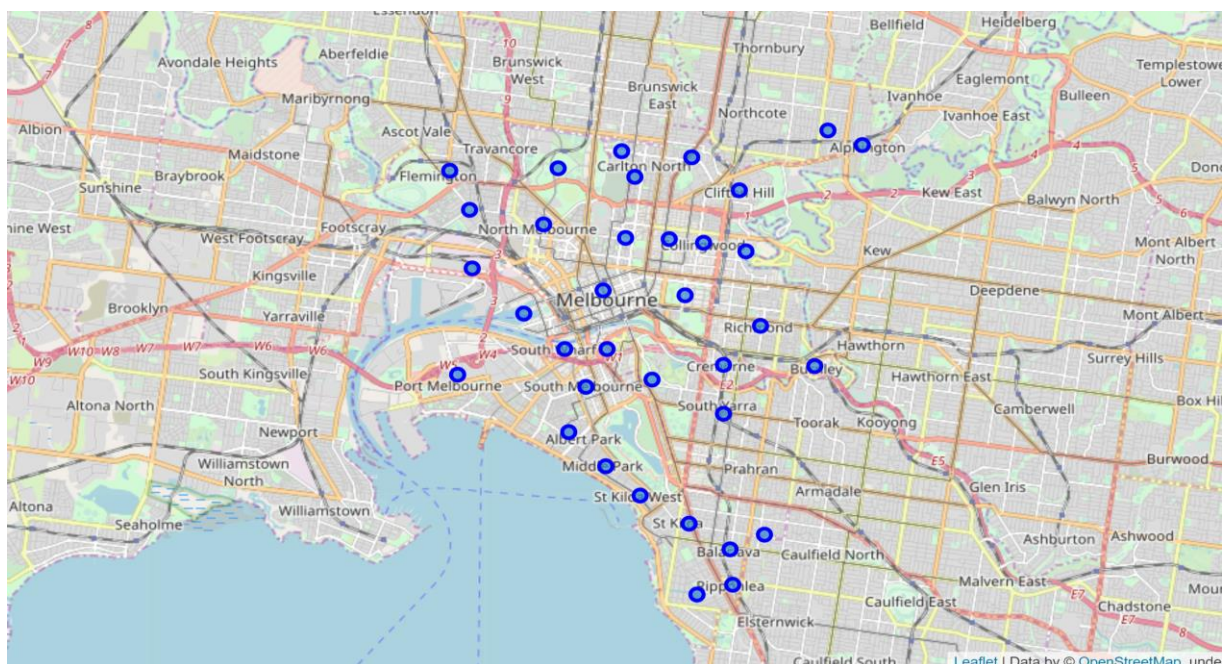
As mentioned above to extract venues for the suburbs Foursquare Places API has been used. There are many endpoints available on Foursquare to get various GET requests. There are a total of 10 categories with unique category id, which can be passed in the URL to extract venues of only that particular category.

```
4d4b7104d754a06370d81259 Arts & Entertainment
4d4b7105d754a06372d81259 College & University
4d4b7105d754a06373d81259 Event
4d4b7105d754a06374d81259 Food
4d4b7105d754a06376d81259 Nightlife Spot
4d4b7105d754a06377d81259 Outdoors & Recreation
4d4b7105d754a06375d81259 Professional & Other Places
4e67e38e036454776db1fb3a Residence
4d4b7105d754a06378d81259 Shop & Service
4d4b7105d754a06379d81259 Travel & Transport
```

A function has been defined to get all nearby food venues for the suburb locations in the dataset. Since our focus is to extract food category, food category id has been used in the URL. After extracting the venues data, only those venue categories which are unique and represent a cuisine are selected. This has been done by removing all the common and irrelevant categories from the dataset.

Exploratory Data Analysis

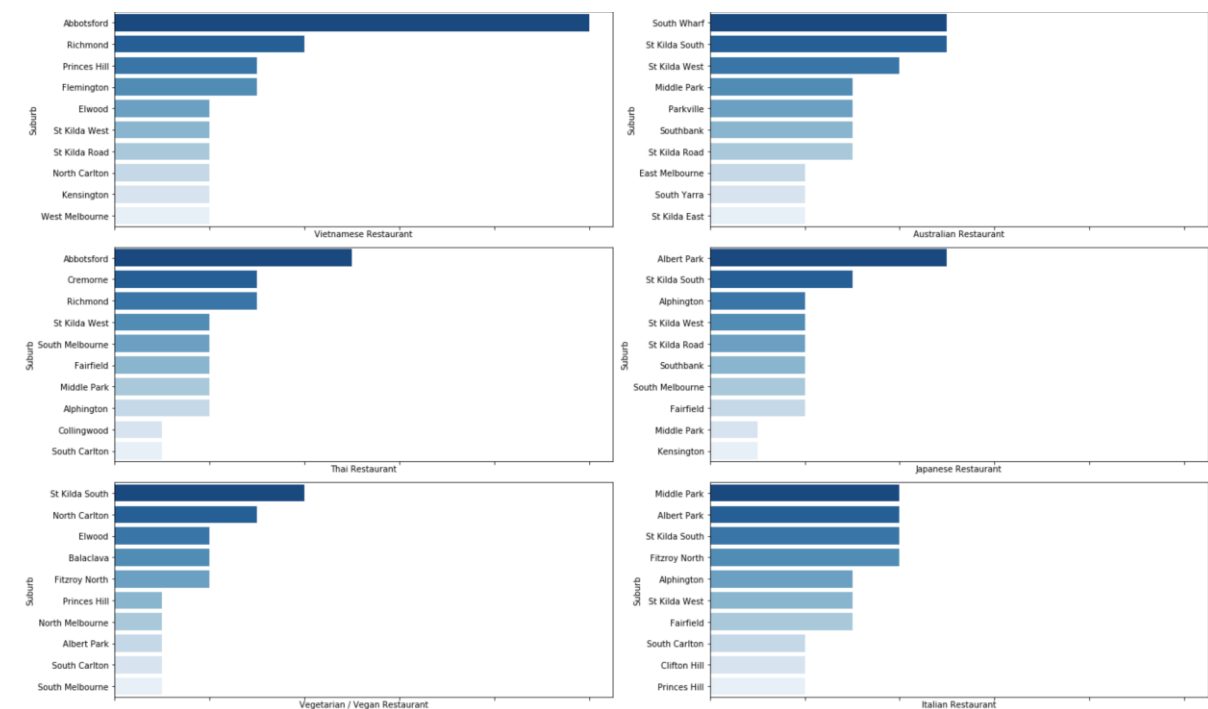
A folium map of Melbourne is created with all suburbs superimposed on it to get an initial idea of where are the suburbs situated and who are their neighbours. The dataset which has 3 cities and 35 suburbs is used for this purpose.

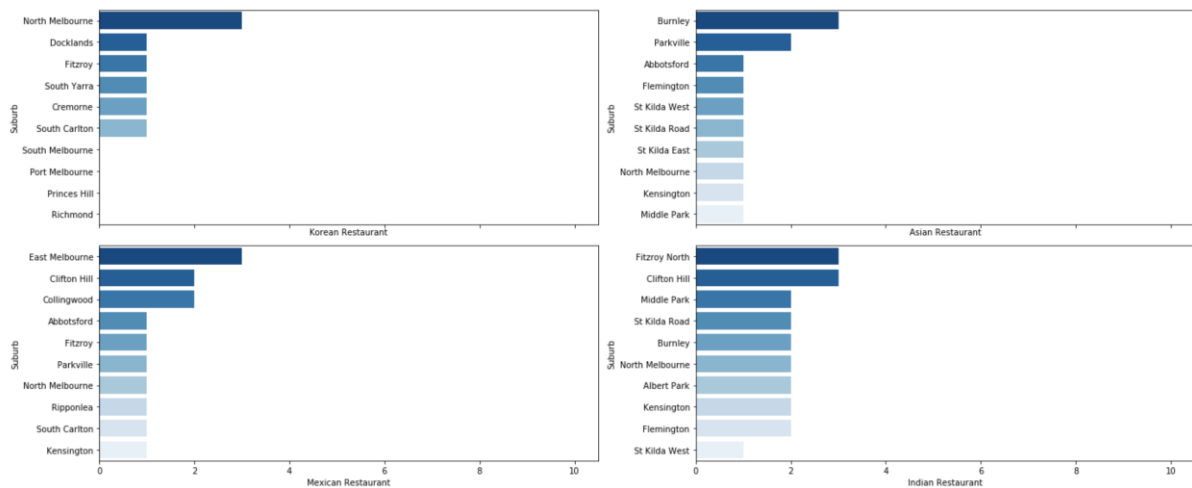


After the final dataset is prepared. Top 10 venue categories are calculated by counting the occurrence frequency of the type of restaurants in the data. Below image shows are Vietnamese, Australian and Thai restaurants are the top occurring restaurants.

	count	mean	std	min	25%	50%	75%	max
Vietnamese Restaurant	35.0	1.228571	1.848461	0.0	0.0	1.0	2.0	10.0
Australian Restaurant	35.0	1.342857	1.454058	0.0	0.0	1.0	2.0	5.0
Thai Restaurant	35.0	0.971429	1.124218	0.0	0.0	1.0	1.0	5.0
Japanese Restaurant	35.0	0.942857	1.083102	0.0	0.0	1.0	1.0	5.0
Vegetarian / Vegan Restaurant	35.0	0.685714	0.963188	0.0	0.0	0.0	1.0	4.0
Italian Restaurant	35.0	1.428571	1.266903	0.0	1.0	1.0	2.0	4.0
Korean Restaurant	35.0	0.228571	0.598317	0.0	0.0	0.0	0.0	3.0
Asian Restaurant	35.0	0.428571	0.698137	0.0	0.0	0.0	1.0	3.0
Mexican Restaurant	35.0	0.514286	0.742469	0.0	0.0	0.0	1.0	3.0
Indian Restaurant	35.0	0.857143	0.943799	0.0	0.0	1.0	1.5	3.0

To further deep dive of the data below visualization is presented, which shows the corresponding top suburbs for each of the above top 10 types of restaurants.





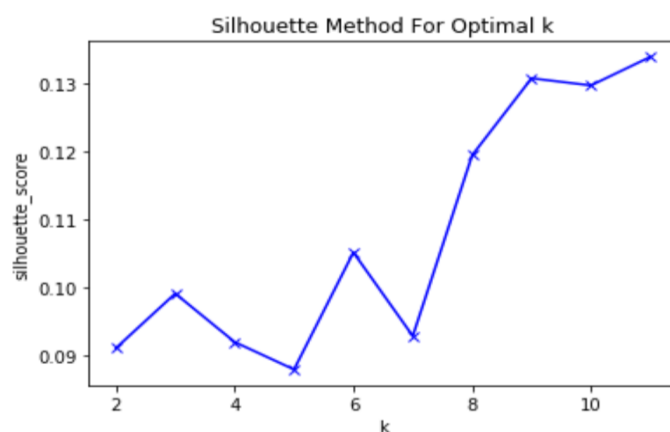
Since a suburb's cuisine preference can be defined by top 5 venues in its vicinity, below table is prepared. It shows top 5 most common venues for each Suburb.

	Suburb	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Abbotsford	Vietnamese Restaurant	Thai Restaurant	Chinese Restaurant	Ramen Restaurant	Asian Restaurant
1	Albert Park	Japanese Restaurant	Italian Restaurant	Indian Restaurant	Vegetarian / Vegan Restaurant	Thai Restaurant
2	Alphington	Italian Restaurant	Thai Restaurant	Japanese Restaurant	Greek Restaurant	Lebanese Restaurant
3	Balaclava	Vegetarian / Vegan Restaurant	Turkish Restaurant	Sushi Restaurant	Malay Restaurant	Italian Restaurant
4	Burnley	Asian Restaurant	Indian Restaurant	Australian Restaurant	Sushi Restaurant	Vietnamese Restaurant

Machine Learning

For the clustering analysis k-means algorithm has been used. K-means is especially useful to discover insights from unlabelled data. Choosing optimal k-value is very important to do a meaning analysis.

To determine the optimal k-value, Silhouette method has been used. As per the graph, k value is optimal at 9. Other k values such as 6 and 8 are also good. However, large k value means large number of resulting clusters. Having huge number of clusters makes it difficult to correctly interpret them. Anything more than 5 or 6 is considered too many clusters to interpret. Therefore, it is better to choose a smaller k value which performs better in a relative sense.

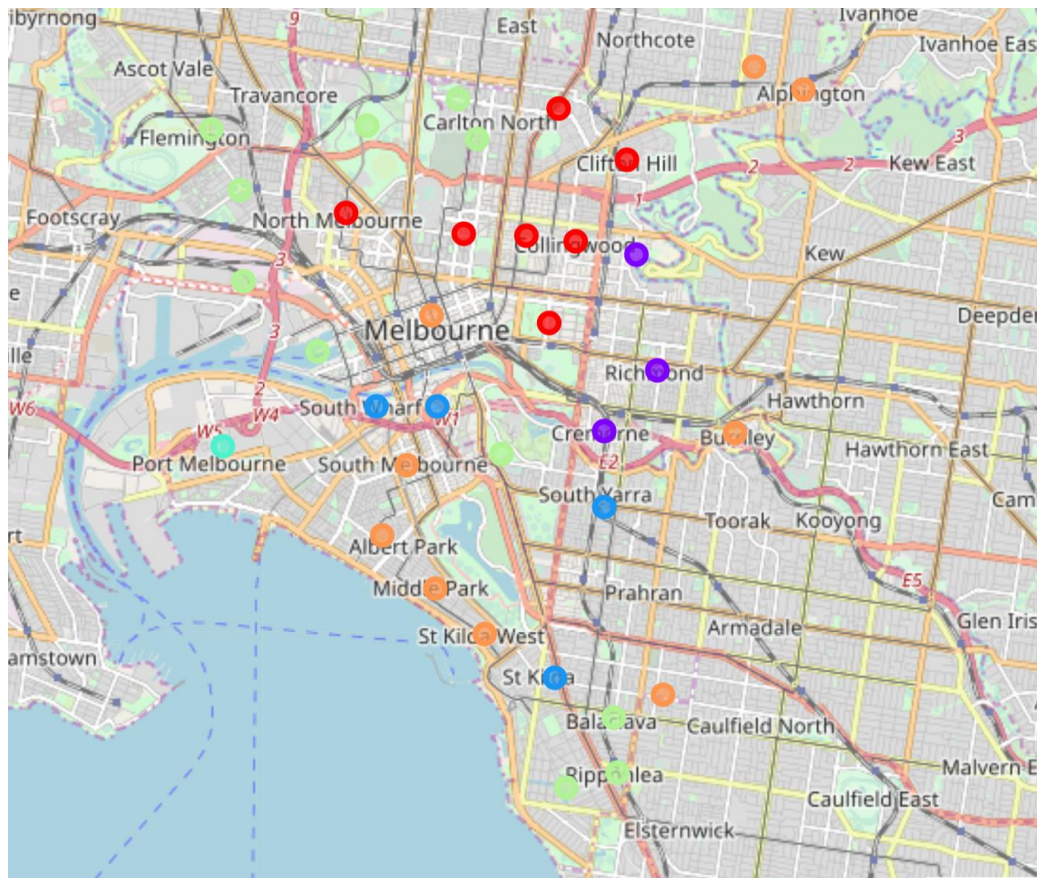


Let's choose $k = 4, 6$ and 8 and look at the results.

For $k = 4$. There is no clear distinction between resultant cluster segments. Every cluster seems similar to the other. All types of restaurants are figured in every cluster. This makes no impact in the analysis. Therefore, $k = 4$ is ignored.

For $k = 8$, there are too many clusters with few members in them. Although on an overall level it seems fine, it is difficult to interpret clusters with just 1 or 2 suburbs in them. Therefore, $k = 8$ is ignored.

For this analysis, optimal k value is 6 . Clusters look meaningful and manageable at the same time when $k = 6$. There is equal division of dominance and mixed clusters. Few clusters are dominated by 1 cuisine whereas others are offering 2 or 3 major cuisines. Below map shows the suburb clusters when $k = 6$.



Results Section

Cluster 0: Majorly there are Indian, Mexican and Italian restaurants. We can see that Italian is the 1st common venue in South Carlton which is a known observation.

cluster_0						
	Suburb	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
6	Clifton Hill	Indian Restaurant	Italian Restaurant	Mexican Restaurant	Moroccan Restaurant	Mediterranean Restaurant
7	Collingwood	Mexican Restaurant	Ramen Restaurant	Sushi Restaurant	Chinese Restaurant	French Restaurant
10	East Melbourne	Mexican Restaurant	Australian Restaurant	Italian Restaurant	Argentinian Restaurant	Thai Restaurant
13	Fitzroy	Korean Restaurant	Australian Restaurant	Indian Restaurant	Mexican Restaurant	Brazilian Restaurant
14	Fitzroy North	Italian Restaurant	Indian Restaurant	Vegetarian / Vegan Restaurant	Mexican Restaurant	Turkish Restaurant
19	North Melbourne	Korean Restaurant	Indian Restaurant	Italian Restaurant	Persian Restaurant	Vegetarian / Vegan Restaurant
25	South Carlton	Italian Restaurant	Korean Restaurant	Australian Restaurant	Persian Restaurant	Indian Restaurant

Cluster 1: Asian Cluster. All restaurants in this cluster and areas are offering Asian cuisines. It has Vietnamese, Thai, Chinese, Sushi etc.

cluster_1						
	Suburb	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Abbotsford	Vietnamese Restaurant	Thai Restaurant	Chinese Restaurant	Ramen Restaurant	Asian Restaurant
8	Cremorne	Thai Restaurant	Sushi Restaurant	Korean Restaurant	Vegetarian / Vegan Restaurant	Italian Restaurant
23	Richmond	Vietnamese Restaurant	Thai Restaurant	Ramen Restaurant	Dumpling Restaurant	Greek Restaurant

Cluster 2: Australian Cluster. This cluster is all about local Australian restaurants as it is the same 1st most common venue for all 4 locations

cluster_2						
	Suburb	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
27	South Wharf	Australian Restaurant	Japanese Restaurant	American Restaurant	Italian Restaurant	Caucasian Restaurant
28	South Yarra	Australian Restaurant	Korean Restaurant	Szechuan Restaurant	Vietnamese Restaurant	French Restaurant
29	Southbank	Australian Restaurant	Japanese Restaurant	Italian Restaurant	American Restaurant	Thai Restaurant
33	St Kilda South	Australian Restaurant	Italian Restaurant	Vegetarian / Vegan Restaurant	Japanese Restaurant	Taco Place

Cluster 3: Port Melbourne Cluster. This cluster has only 1 location with Japanese and Mexican offerings

cluster_3						
	Suburb	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
21	Port Melbourne	Japanese Restaurant	Mexican Restaurant	Chinese Restaurant	Eastern European Restaurant	Italian Restaurant

Cluster 4: This cluster is Vegetarian/Vegan and Vietnamese cuisines along with few Middle Eastern and Turkish Restaurants

cluster_4						
	Suburb	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
3	Balaclava	Vegetarian / Vegan Restaurant	Turkish Restaurant	Sushi Restaurant	Malay Restaurant	Italian Restaurant
9	Docklands	Middle Eastern Restaurant	Korean Restaurant	Southern / Soul Food Restaurant	Indonesian Restaurant	Italian Restaurant
11	Elwood	Vegetarian / Vegan Restaurant	Vietnamese Restaurant	Turkish Restaurant	Thai Restaurant	Kebab Restaurant
15	Flemington	Vietnamese Restaurant	Indian Restaurant	Malay Restaurant	Yemeni Restaurant	Noodle House
16	Kensington	Indian Restaurant	Vietnamese Restaurant	Malay Restaurant	Yemeni Restaurant	Noodle House
18	North Carlton	Vegetarian / Vegan Restaurant	Italian Restaurant	Vietnamese Restaurant	Brazilian Restaurant	Japanese Restaurant
20	Parkville	Australian Restaurant	Asian Restaurant	Kebab Restaurant	Dumpling Restaurant	Italian Restaurant
22	Princes Hill	Vietnamese Restaurant	Italian Restaurant	Middle Eastern Restaurant	Lebanese Restaurant	Vegetarian / Vegan Restaurant
24	Ripponlea	Australian Restaurant	Turkish Restaurant	Taco Place	Falafel Restaurant	Mexican Restaurant
31	St Kilda Road	Australian Restaurant	Sushi Restaurant	Indian Restaurant	Japanese Restaurant	French Restaurant
34	West Melbourne	Vietnamese Restaurant	Chinese Restaurant	Portuguese Restaurant	Souvlaki Shop	Greek Restaurant

Cluster 5: This cluster has Italian and Japanese restaurants along with few Australian Restaurants

cluster_5						
	Suburb	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
1	Albert Park	Japanese Restaurant	Italian Restaurant	Indian Restaurant	Vegetarian / Vegan Restaurant	Thai Restaurant
2	Alphington	Italian Restaurant	Thai Restaurant	Japanese Restaurant	Greek Restaurant	Lebanese Restaurant
4	Burnley	Asian Restaurant	Indian Restaurant	Australian Restaurant	Sushi Restaurant	Vietnamese Restaurant
5	CBD	Ramen Restaurant	Australian Restaurant	Vietnamese Restaurant	Greek Restaurant	Persian Restaurant
12	Fairfield	Italian Restaurant	Japanese Restaurant	Thai Restaurant	Greek Restaurant	Australian Restaurant
17	Middle Park	Italian Restaurant	Australian Restaurant	Thai Restaurant	Indian Restaurant	Japanese Restaurant
26	South Melbourne	Japanese Restaurant	Indonesian Restaurant	Thai Restaurant	Caucasian Restaurant	Vegetarian / Vegan Restaurant
30	St Kilda East	Australian Restaurant	Sushi Restaurant	African Restaurant	Vietnamese Restaurant	German Restaurant
32	St Kilda West	Australian Restaurant	Italian Restaurant	Japanese Restaurant	Thai Restaurant	Vietnamese Restaurant

Discussion Section

From results we can see that there are 3 specific clusters and 3 clusters are a mix of 2 or 3 major cuisines

Cluster Number	Cuisine/Restaurants
0	Indian, Mexican and Italian
1	Asian
2	Australian
3	Japanese and Mexican
4	Vegetarian/Vegan, Vietnamese and Middle Eastern
5	Italian, Japanese and Australian

General Observations

From the above table we can see that it represents most regions of the world. There is Asian, Mexican, Middle Eastern, Italian and Australian. Therefore, it is safe to say that there is multi-cultural food available in Melbourne.

It also reflects the migration trends. There are significant number of Vietnamese, Japanese and Italian restaurants which coincide with the information from Melbourne city website.

However, one interesting observation is no dominance of Chinese cuisine in the top 10 and also in the most common venues. The underlying reason for this could be, Chinese restaurants go with specific names like Sushi restaurant, Dim Sum restaurant, Asian restaurant etc. which leads to less venues per each category thus not showing a strong presence.

Recommendations for Businesses

Cluster 1 i.e. Abbotsford, Cremorne and Richmond suburbs are covered with Asian restaurants making it saturated. Therefore, it might not be a good area of choice for starting a new Asian restaurant if a business wants to distinguish itself from other players

Cluster 4 i.e. In Balaclava, Elwood and North Carlton 1st common venue is Vegetarian/Vegan restaurant. With the rise in veganism trend, this could be an interesting business opportunity to start a vegan restaurant in areas surrounding these suburbs where there seems to be demand.

Conclusion Section

From all the facts and figures, Melbourne is truly a food multicultural city. It has vast cuisine offerings from across the countries and regions. The analysis also reflects that food scene is going hand in hand with the earlier and current migration trends.

References

<https://www.melbourne.vic.gov.au/about-melbourne/research-and-statistics/city-population/Pages/residents-profiles.aspx>