Capstone Project - Week 4-5

Vishnu Vinayamohanan

July 28, 2021

1   Introduction

Australia is a multi-cultural country which has got more than 200 nationalities living in the country. The diversity in population brought numerous cuisines from all over the world to Australia. The people of Australia are very keen to explore different cusine from across the globe. The aim of this project is to analyze which city in Australia is best suited for starting a new restaurant business.

1.1   Problem Statement

Analyzing which cities is Australia is best suited for starting Restaurants.

2   Data Collection

Data used in this project are as follows;

2.1   Venue Data

The venue data has been extracted using the Foursquare API. This data contains venue recommendations for all cities in Australia and is used to study the popular venues of different neighborhoods.

2.2   Australian Cities

Main Australian cities are obtained from the url "https://www.latlong.net/category/cities-14-15.html". The data sets consists of the City names, latitude, and longitude information for each neighborhood. The data will be cleaned to remove redundancy and difference in the State Names as explained in the notebook. The data is read into a pandas data frame using the read_html() method.

2.3   Coordinates

The geographical coordinates for Australia data has been obtained from the GeoPy library in python.

3   Methodology

- Details of the Australian Cities would be scrapped from the url mentioned in section 2.2

- The coordinates of each of these Australian Cities will be fetched using the python Geopy package.

- Foursquare API search feature would be enabled to collect the nearby venues of the cities. Due to http request limitations the number of places per neighborhood parameter would reasonably be set to 200 and the radius parameter would be set to 1000.

- Folium- Python visualization library would be used to visualize the Australian cities cluster distribution.

- Unsupervised machine learning algorithm K-mean clustering would be applied to form the clusters of Australian cities.

4 Discussion

4.1 Fetching Main cities of Australia

The main cities of Australia are fetched from the url "https://www.latlong.net/category/cities-14-15.html". The corresponding python script are shown in Fig. 1. Some states like QLD is represented as Queensland and Tasmania as TAS. The data frame is cleaned to make the state names which are represented as different names to a common name.

```
[5]  aus_cities = pd.read_html('https://www.latlong.net/category/cities-14-15.html')[-1]
     aus_cities['Cities'] = aus_cities['Place Name'].apply(lambda x: x.split(',')[0])
     aus_cities['State'] = aus_cities['Place Name'].apply(lambda x: x.split(',')[1])
     aus_cities.drop(['Place Name', 'Latitude', 'Longitude' ], axis=1, inplace=True)
     aus_cities.head(5)
```

	Cities	State
0	Tamworth	NSW
1	Queanbeyan	NSW
2	Penrith	NSW
3	Newcastle	NSW
4	Liverpool	NSW

Figure 1: Fetching Main cities of Australia

4.2 Coordinates of the cities using Geopy

Once the data is cleaned, the latitude and longitude of the cities are obtained using the python Geopy package. The corresponding results are shown in Fig.2

	Cities	State	Latitude	Longitude
0	Tamworth	NSW	-31.0906	150.929
1	Queanbeyan	NSW	-35.355	149.233
2	Penrith	NSW	-33.7537	150.698
3	Newcastle	NSW	-32.9281	151.773
4	Liverpool	NSW	-33.9224	150.923
5	Lithgow	NSW	-33.4828	150.153
6	Goulburn	NSW	-34.757	149.716
7	Dubbo	NSW	-32.247	148.601
8	Cessnock NSW	QLD	-32.8329	151.355
9	Campbelltown	NSW	-34.8809	138.662

Figure 2: Fetching Coordinates of the cities using Geopy

4.3 Visualize the cities using folium

Once the coordinates are obtained using the Geopy, the cities are visulaized using the python folium package. The corresponding map is shown in Fig.3

Figure 3: Map of Australian Cities

4.4 Obtaining the venue details in each cities

The next step is to fins the venue details in each cities using the Foursquare API. The corresponding results are shown in Fig.4

	Cities	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Adelaide	100	100	100	100	100	100
1	Albany	26	26	26	26	26	26
2	Albury	39	39	39	39	39	39
3	Ararat	14	14	14	14	14	14
4	Armadale	21	21	21	21	21	21
...	...	...	...	...	...	...	...
81	Wanneroo	14	14	14	14	14	14
82	Warrnambool	29	29	29	29	29	29
83	Westmead	9	9	9	9	9	9
84	Whyalla	6	6	6	6	6	6
85	Wollongong	82	82	82	82	82	82

86 rows × 7 columns

Figure 4: Venues in each cities obtained using Foursquare API

4.5 One-hot encoding

The next step is to applying One-hot Encoding to see which categories belong in which cities.The corresponding results are shown in Fig.5

	Cities	African Restaurant	American Restaurant	Aquarium	Arcade	Argentinian Restaurant	Art Gallery	Art Museum	Arts & Crafts Store	Arts & Entertainment	Asian Restaurant	Athletics & Sports	Australian Restaurant	Auto Garage	BBQ Joint	Bagel Shop	Bakery	Bar	Beach	Bed & Breakfast	Beer Bar	Beer Garden	Big Box Store	Bistro	Boat Launch	Boat or Ferry	Bookstore	Botanical Garden	Bouti
0	Tamworth	0.04	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.000000	0.000	0.000000	0.0	0.0	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.0	0.040000	0.0	0.0	0.0	0.000000	0.0
1	Queanbeyan	0.00	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.000000	0.000	0.000000	0.0	0.0	0.000000	0.062500	0.000000	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.000000	0.0
2	Penrith	0.00	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.014493	0.000	0.000000	0.0	0.0	0.000000	0.000000	0.014493	0.0	0.0	0.0	0.0	0.014493	0.0	0.0	0.0	0.000000	0.0
3	Newcastle	0.00	0.0	0.0	0.0	0.0	0.010638	0.0	0.0	0.0	0.000000	0.000	0.010638	0.0	0.0	0.010638	0.000000	0.085106	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.010638	0.0
4	Liverpool	0.00	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.000000	0.000	0.000000	0.0	0.0	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.000000	0.0
5	Lithgow	0.00	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.000000	0.000	0.000000	0.0	0.0	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.000000	0.0
6	Goulburn	0.00	0.0	0.0	0.0	0.0	0.045455	0.0	0.0	0.0	0.000000	0.000	0.000000	0.0	0.0	0.000000	0.045455	0.045455	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.000000	0.0
7	Dubbo	0.00	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.000000	0.000	0.027027	0.0	0.0	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.000000	0.0
8	Cessnock NSW	0.00	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.000000	0.000	0.000000	0.0	0.0	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.0	0.045455	0.0	0.0	0.0	0.000000	0.0
9	Campbelltown	0.00	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.000000	0.125	0.000000	0.0	0.0	0.000000	0.125000	0.000000	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.000000	0.0

10 rows × 256 columns

Figure 5: One hot encoding results

4.6 Finding top ten venues in cities

In order to find the popular business in each cities, the top ten venues are calculated. This will give us an over view about the popular business in each cities. The corresponding results are shown in Fig.6

	Cities	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Tamworth	Supermarket	Fast Food Restaurant	Department Store	Café	African Restaurant	Electronics Store	Steakhouse	Shopping Mall	Noodle House	Gas Station
1	Queanbeyan	Café	Pub	Park	Supermarket	Thai Restaurant	Hotel	Performing Arts Venue	Sandwich Place	Bakery	Pizza Place
2	Penrith	Café	Fast Food Restaurant	Thai Restaurant	Electronics Store	Supermarket	Department Store	Performing Arts Venue	Sandwich Place	Steakhouse	Shopping Mall
3	Newcastle	Café	Bar	Pub	Restaurant	Vietnamese Restaurant	Sandwich Place	Pizza Place	Breakfast Spot	Supermarket	Japanese Restaurant
4	Liverpool	Café	Fast Food Restaurant	Gym / Fitness Center	Sandwich Place	Pharmacy	Department Store	Supermarket	Shopping Mall	Lebanese Restaurant	Coffee Shop
...	...	...	...	...	...	...	...	...	...	...	...
81	Gladstone QLD	Multiplex	Convenience Store	Restaurant	Coffee Shop	Asian Restaurant	Pizza Place	Indian Restaurant	Sandwich Place	Football Stadium	Food Truck
82	Gosford	Pizza Place	Miscellaneous Shop	Café	Playground	Platform	Fast Food Restaurant	Sandwich Place	Pharmacy	Food	Football Stadium
83	Mount Gambier	Café	Fast Food Restaurant	Tourist Information Center	Pub	Hotel	Food & Drink Shop	Shopping Mall	Sandwich Place	Supermarket	Cave
84	Sydney	Café	Coffee Shop	Speakeasy	Hotel	Cocktail Bar	Bar	Shopping Mall	Japanese Restaurant	Hobby Shop	Restaurant
85	Glenore Grove	Convenience Store	Zoo	Falafel Restaurant	Frame Store	Fountain	Football Stadium	Food Truck	Food Court	Food & Drink Shop	Food

86 rows × 11 columns

Figure 6: Top ten venues in each cities

4.7 Clustering cities

The next step is to cluster the cities using K-mean unsupervised technique. For finding the best cluster length, the Silhouette Score is calculated and the corresponding result is shown in Fig.7. Based on the score, the cluster length of 8 is chosen for further analysis.
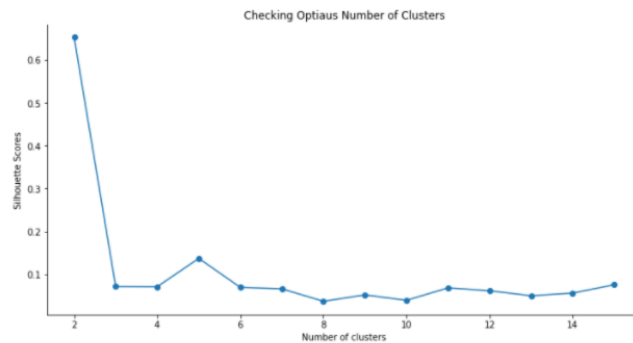
Figure 7: Silhouette Score

once the cluster length is finalized, a folium map is created which shows the cities in each cluster and their top ten venues. The corresponding results are shown in Fig.8
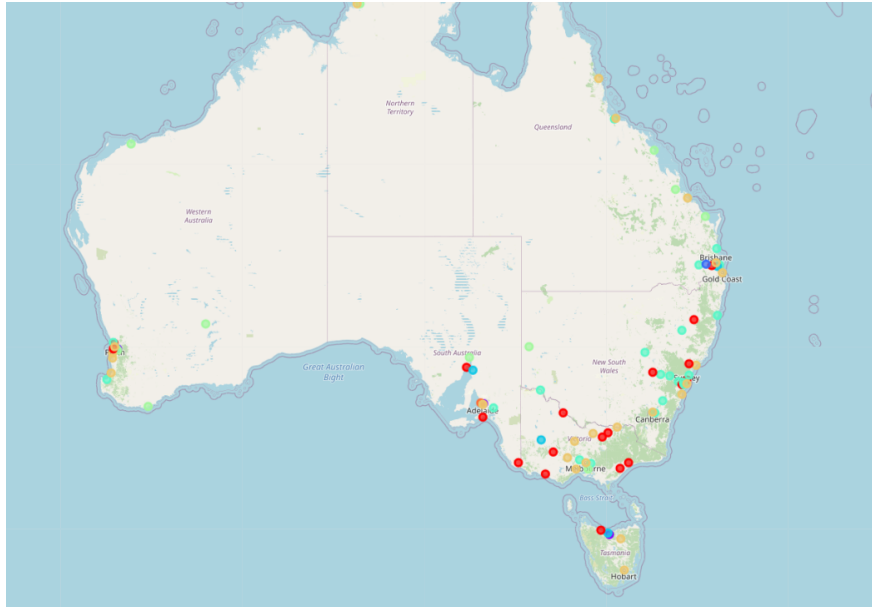

Figure 8: Map showing cities in each cluster and their top ten venues

4.8 Finding best cluster for restaurant

To find the best cluster suited for restaurant business, 5 clusters are analyzed and the corresponding results are shown in Fig.9 - Fig.13

▼ Cluster 1

```
aus_merged.loc[aus_merged['Cluster Labels'] == 0, aus_merged.columns[[0] + [1] + list(range(5, aus_merged.shape[1]))]]
```

	Cities	State	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
4	Liverpool NSW		Café	Fast Food Restaurant	Gym / Fitness Center	Sandwich Place	Pharmacy	Department Store	Supermarket	Shopping Mall	Lebanese Restaurant	Coffee Shop
8	Cessnock NSW	QLD	Café	Pub	Fast Food Restaurant	Supermarket	Hotel	Grocery Store	Department Store	Hotel Bar	Discount Store	Coffee Shop
13	Armidale	NSW	Café	Pub	Fast Food Restaurant	Park	Motel	Supermarket	Sandwich Place	Hotel	Gym / Fitness Center	Gas Station
14	Wanneroo	WA	Fast Food Restaurant	Café	Portuguese Restaurant	Restaurant	Department Store	Shopping Mall	Pet Store	Liquor Store	Coffee Shop	Steakhouse
25	Fremantle	WA	Café	Japanese Restaurant	Supermarket	Hostel	Tapas Restaurant	Gas Station	Breakfast Spot	South American Restaurant	Seafood Restaurant	Sandwich Place
29	Warrnambool	VIC	Café	Hotel	Grocery Store	Pub	Multiplex	South American Restaurant	Fast Food Restaurant	Supermarket	Noodle House	Gas Station
30	Wangaratta	VIC	Department Store	Fast Food Restaurant	Café	Sandwich Place	Australian Restaurant	Hotel Bar	Pizza Place	Performing Arts Venue	Farmers Market	Fish & Chips Shop
31	Swan Hill	VIC	Café	Fast Food Restaurant	Supermarket	Auto Garage	Hotel	Theme Park	RV Park	Park	Thai Restaurant	Gas Station
32	Sale	VIC	Café	Fast Food Restaurant	Hotel	Grocery Store	Discount Store	Electronics Store	Restaurant	Chinese Restaurant	Sandwich Place	Shopping Mall
40	Benalla	VIC	Café	Bakery	Train Station	Fast Food Restaurant	Pub	Australian Restaurant	Supermarket	Fish & Chips Shop	Fishing Store	Flea Market
41	Bairnsdale	VIC	Grocery Store	Fast Food Restaurant	Café	Supermarket	Gas Station	Sandwich Place	Construction & Landscaping	Convenience Store	Hotel	Department Store
42	Ararat	VIC	Café	Fast Food Restaurant	Bakery	Grocery Store	Discount Store	Food	Train Station	Italian Restaurant	Park	Tourist Information Center
44	Burnie	TAS	Fast Food Restaurant	Bar	Gas Station	Sandwich Place	Bakery	Seafood Restaurant	Electronics Store	Grocery Store	Rental Car Location	Café
46	Whyalla	SA	Café	Hotel Bar	Department Store	Motel	Discount Store	Scenic Lookout	Zoo	Fast Food Restaurant	Fish & Chips Shop	Fishing Store
47	Victor Harbor	SA	Café	Supermarket	Hotel	Train Station	Convenience Store	Park	Fast Food Restaurant	Sandwich Place	Campground	Museum
73	Orange	NSW	Café	Supermarket	Department Store	Theater	Park	Sports Club	Australian Restaurant	Train Station	Coffee Shop	Pub
77	Ipswich	QLD	Café	Fast Food Restaurant	Hotel	Bar	Portuguese Restaurant	Performing Arts Venue	Park	Coffee Shop	Sandwich Place	Nature Preserve
80	Terrey Hills	NSW	Café	Indian Restaurant	Pub	Golf Course	Fruit & Vegetable Store	Bus Station	Flower Shop	Zoo	Fish & Chips Shop	Fishing Store
86	Mount Gambier	SA	Café	Fast Food Restaurant	Tourist Information Center	Pub	Hotel	Food & Drink Shop	Shopping Mall	Sandwich Place	Supermarket	Cave

Figure 9: Cluster-1

▼ Cluster 2

```
[46] aus_merged.loc[aus_merged['Cluster Labels'] == 1, aus_merged.columns[[0] + [1] + list(range(5, aus_merged.shape[1]))]]
```

	Cities	State	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
9	Campbelltown	NSW	Grocery Store	Bakery	Trail	Athletics & Sports	Café	Chinese Restaurant	Flower Shop	Fish & Chips Shop	Fishing Store	Flea Market
15	Gosnells	WA	Grocery Store	Fast Food Restaurant	Train Station	Sandwich Place	Fishing Store	Event Space	Falafel Restaurant	Farmers Market	Fish & Chips Shop	Zoo
16	Spearwood	WA	Event Service	Thai Restaurant	Shopping Mall	Fish & Chips Shop	Grocery Store	Zoo	Fishing Store	Farmers Market	Fast Food Restaurant	Flea Market
35	Latrobe City	VIC	Grocery Store	Coffee Shop	Health & Beauty Service	Gift Shop	Flea Market	Farmers Market	Fast Food Restaurant	Fish & Chips Shop	Fishing Store	Zoo

Figure 10: Cluster-2

```
aus_merged.loc[aus_merged['Cluster Labels'] == 2, aus_merged.columns[[0] + [1] + list(range(5, aus_merged.shape[1]))]]
```

	Cities	State	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
88	Glenore Grove	QLD	Convenience Store	Zoo	Falafel Restaurant	Frame Store	Fountain	Football Stadium	Food Truck	Food Court	Food & Drink Shop	Food

Figure 11: Cluster-3

▼ Cluster 4

```
[48] aus_merged.loc[aus_merged['Cluster Labels'] == 3, aus_merged.columns[[0] + [1] + list(range(5, aus_merged.shape[1]))]]
```

	Cities	State	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
36	Horsham	VIC	Supermarket	Fast Food Restaurant	Bakery	Mexican Restaurant	Train Station	Café	Bar	Department Store	Food Court	Food & Drink Shop
43	Devonport	TAS	Fast Food Restaurant	Supermarket	Department Store	Bowling Alley	Grocery Store	Food	Fish & Chips Shop	Fishing Store	Flea Market	Flower Shop
48	Port Pirie	SA	Gas Station	Fast Food Restaurant	Tourist Information Center	Hostel	Supermarket	Flower Shop	Farmers Market	Fish & Chips Shop	Fishing Store	Flea Market
57	Logan City	QLD	Supermarket	Fast Food Restaurant	Grocery Store	Fruit & Vegetable Store	Gym	Performing Arts Venue	Australian Restaurant	Furniture / Home Store	Train Station	Pharmacy

Figure 12: Cluster-4

▼ Cluster 5

```
[49] aus_merged.loc[aus_merged['Cluster Labels'] == 4, aus_merged.columns[[0] + [1] + list(range(5, aus_merged.shape[1]))]]
```

	Cities	State	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Tamworth	NSW	Supermarket	Fast Food Restaurant	Department Store	Café	African Restaurant	Electronics Store	Steakhouse	Shopping Mall	Noodle House	Gas Station
1	Queanbeyan	NSW	Café	Pub	Park	Supermarket	Thai Restaurant	Hotel	Performing Arts Venue	Sandwich Place	Bakery	Pizza Place
2	Penrith	NSW	Café	Fast Food Restaurant	Thai Restaurant	Electronics Store	Supermarket	Department Store	Performing Arts Venue	Sandwich Place	Steakhouse	Shopping Mall
5	Lithgow	NSW	Supermarket	Café	Park	Sandwich Place	Zoo	Farmers Market	Fast Food Restaurant	Fish & Chips Shop	Fishing Store	Flower Shop
6	Goulburn	NSW	Supermarket	Café	Department Store	Social Club	Pub	Thai Restaurant	Bar	Bakery	Sandwich Place	Train Station
7	Dubbo	NSW	Pizza Place	Café	Coffee Shop	Fast Food Restaurant	Supermarket	Motel	Social Club	Food Court	Shopping Mall	
12	Bathurst	NSW	Café	Supermarket	Department Store	Shopping Mall	Fast Food Restaurant	Liquor Store	Clothing Store	Sandwich Place	Brewery	Irish Pub
17	Cannington	WA	Fast Food Restaurant	Department Store	Supermarket	Electronics Store	Clothing Store	Café	Chinese Restaurant	Pool	Pool Hall	Park
18	Bayswater	WA	Malay Restaurant	Vietnamese Restaurant	Supermarket	Café	Thai Restaurant	Gas Station	Dumpling Restaurant	Sandwich Place	Steakhouse	Coffee Shop
24	Joondalup	WA	Pizza Place	Park	Playground	Fast Food Restaurant	Thai Restaurant	Sandwich Place	Mexican Restaurant	Seafood Restaurant	Bakery	Paper / Office Supplies Store
26	Busselton	WA	Fast Food Restaurant	Café	Supermarket	Thai Restaurant	Beer Garden	Tapas Restaurant	Furniture / Home Store	Beach	Pub	Sandwich Place
34	Melton	VIC	Gym	Pub	Fast Food Restaurant	Convenience Store	Sandwich Place	Soccer Field	Bar	Diner	Mexican Restaurant	Café
51	Murray Bridge	SA	Bakery	Shopping Mall	Fast Food Restaurant	Train Station	Asian Restaurant	Boat or Ferry	Supermarket	Grocery Store	Outdoor Sculpture	Park
52	Thuringowa	QLD	Supermarket	Café	Fast Food Restaurant	Department Store	Stadium	Electronics Store	Shopping Mall	Chinese Restaurant	Steakhouse	Bowling Alley
53	Redland City	QLD	Business Service	Carpet Store	Park	Supermarket	Food & Drink Shop	Fast Food Restaurant	Fish & Chips Shop	Fishing Store	Flea Market	Zoo
54	Toowoomba	QLD	Café	Music Venue	Fast Food Restaurant	Movie Theater	Steakhouse	Supermarket	Breakfast Spot	Park	Coffee Shop	Pub
61	Sunshine Coast	QLD	Grocery Store	Boat Launch	Home Service	River	Café	Clothing Store	Miscellaneous Shop	Flea Market	Fast Food Restaurant	Fish & Chips Shop
70	Mildura	VIC	Supermarket	Café	Pharmacy	Steakhouse	Coffee Shop	Chinese Restaurant	Plaza	Sports Bar	Brewery	Fast Food Restaurant
72	Coffs Harbour	NSW	Supermarket	Chinese Restaurant	Café	Botanical Garden	Fishing Store	Shopping Mall	Liquor Store	Electronics Store	Department Store	Construction & Landscaping
81	Bankstown NSW	QLD	Vietnamese Restaurant	Café	Department Store	Fast Food Restaurant	Buffet	Bar	Sports Bar	Donut Shop	Convenience Store	Coffee Shop
82	Westmead	NSW	Golf Course	Indian Restaurant	Rental Car Location	Snack Place	Café	Fast Food Restaurant	Sandwich Place	Bar	Distillery	Discount Store
85	Gosford	NSW	Pizza Place	Miscellaneous Shop	Café	Playground	Platform	Fast Food Restaurant	Sandwich Place	Pharmacy	Food	Football Stadium

Figure 13: Cluster-5

5 Results and Discussion

Analysis of the five clusters obtained shows which clusters are more suited for restaurants and hotels. Cities in clusters 2, 3, 4 contain a small percentage of restaurants, hotels, cafe and pubs in their top 10 common venues. These clusters contain a higher degree of other venues like grocery store, convenience store, and supermarkets. Thus, they are not well suited for opening a new restaurant. On the other hand, cities in clusters 1 and 5 contain a much higher degree of restaurants, multiplex, cafes, bars. Thus, it can be concluded that the cities in these clusters would be well suited for opening a new restaurant.

Comparing clusters 1 and 5, cities in cluster 1 seem to be more suited for starting a restaurant since they contains a larger percentage of cafes and restaurants in the top 10 most common venues than cluster 5. Thus, it is recommended that the new restaurant can be opened in the cities belonging to cluster 1.

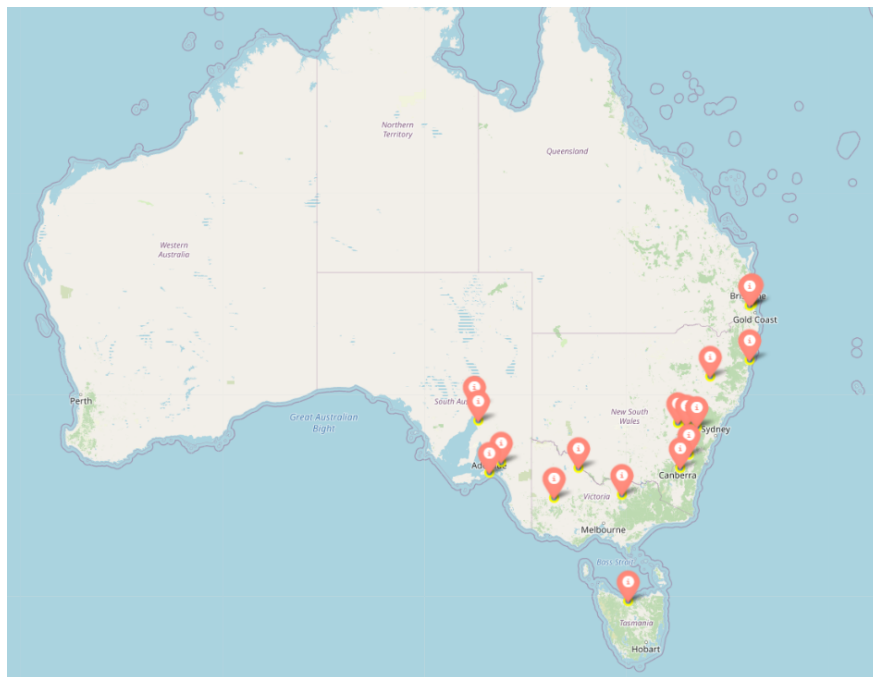The cities in cluster-1 are shown in the below map.

Figure 14: Map showing cities in Cluster-1

6 Conclusion

We have successfully analyzed the main cities in Australia for determining which would be the best city for opening a new restaurant. Our analysis shows that the neighborhoods in cluster 1 are best suited for new restaurant.