

FINDING A LOCATION TO OPEN A COFFEE SHOP IN HOBART USING INCOME AND VENUES DATA FOR GREATER HOBART NEIGHBOURHOODS

1. INTRODUCTION

1.1. Problem Description

It is well known that the location of shop is an important factor in determining its success. In addition to researching the coffee business, competition, defining your vision and planning food and drinks menu as well as decor, it is important to seriously consider the location of your shop. The requirement for a shop location is based on multiple factors including its closeness and diversity of nearby venues, demographics of the neighbourhood, closeness to businesses, schools and hospitals and availability of parking or closeness to transport. People who regularly visit a coffee shop live or work 5-10minutes walk from the location. Other considerations include corner of location and rental rates. My client is interested in opening coffee shop in Hobart and has the following location requirements:

- i. It must be close to other venues but avoid locations already with high concentration of similar businesses i.e. coffee shops. The more diverse a location is the better.
- ii. Location with high population density.
- iii. In a neighbourhood with residents who have time and money to regularly spend visit cafes'

The project is to find a good neighbourhood to open a coffee shop. I will use income and population density to select neighbourhoods. These neighbourhoods will be compared using venues.

1.2. About Hobart, Tasmania, Australia

Hobart is the capital city of Tasmania, a small island state of Australia. It is known for its natural beauty, clean air and waters. Its backdrop is Mount wellington with hiking and cycling trails. It is the smallest state capital with a population of 240,000. The city offers resident and tourists with has great selection of food and wine. Greater Hobart is divided into 6 regions called Local Government Areas. These are the cities of Hobart, Glenorchy and Clarence, as well as municipalities of Kingborough, Brighton, Sorell. Each LGAs has localities called suburbs. Some suburbs near each other share post codes.

2. DATA DESCRIPTION

2.1. Neighbourhoods data: A csv file was acquired from [AustralianTownsList.com](https://australiantownslist.com)

A csv file was acquired from AustralianTownsList.com. The data was last updated April 2020. The file has columns for ID, Name of location, type of location, urban area (the location falls under), state, post code, coordinates, population, median income, and local government area (LGA). We will select locations of type suburbs in LGAs of greater Hobart, in state of Tasmania. For our analysis we will use the following columns – location/suburb, post code, LGA, income, coordinates, population, and area size. A population density column will be added using population and area size columns

2.2. Venues data

I used Foursquare API to get venues in the chosen neighbourhoods.

3. METHODOLOGY

Basic statistical analysis and visualisation of data to identify neighbourhood with high household median income and densely populated. Median value of neighbourhood median income and population density will be used to select neighbourhoods with high/upper middle-income neighbourhoods and high population densities. These neighbourhoods will then be used in K-means clustering algorithm using venues found in the Foursquare API. They will be clustered into similar groups. These groups will then be assessed to find what is unique about each group. A cluster with good diversity of venues and having cafes/coffee shop as one of the top 10 venues will be selected (café_cluster). Among those neighbourhoods we will identify neighbourhoods where cafes/coffee shop is not in top 2 most common venues. These neighbourhoods are fairly like other neighbourhood with cafes as top venues which might indicate that the neighbourhood has room for more coffee shop/cafe.

4. ANALYSIS, VISUALISATION AND MODELING

4.1. The Dataframe

The dataframe has 6 Local Government Areas and 131 neighbourhoods in Greater Hobart region. And columns are as shown in figure 1.

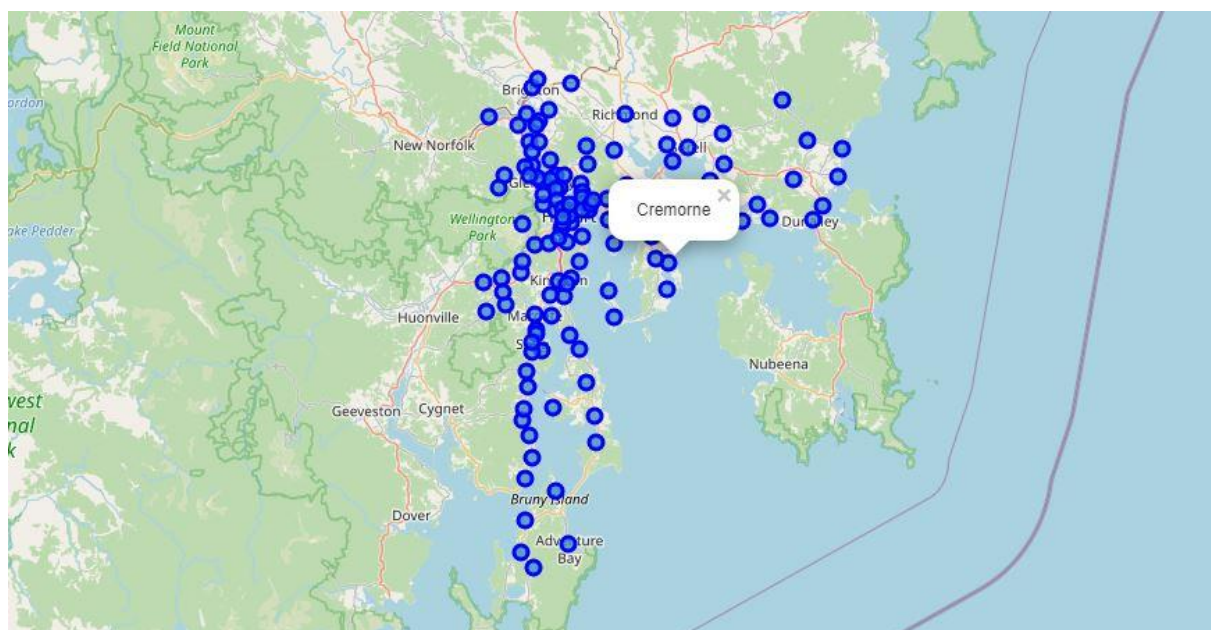
Figure 1: Dataframe for all neighbourhoods in Greater Hobart.

	Neighborhood	postcode	LGA	latitude	longitude	median_income	population	area_sq_km	pop_density
0	Acton Park	7170	Clarence	-42.86574	147.46989	43472.0	2078	19.267	107.852805
1	Adventure Bay	7150	Kingborough	-43.35579	147.32612	22464.0	195	11.427	17.064846
2	Allens Rivulet	7150	Kingborough	-43.01068	147.20406	37544.0	487	11.972	40.678249
3	Alonnah	7150	Kingborough	-43.32148	147.24218	24544.0	137	15.170	9.030982
4	Apollo Bay	7150	Kingborough	-43.15912	147.29609	39416.0	23	2.295	10.021786

4.2 Visualise all Greater Hobart Region neighbourhoods

Greater Hobart region is a wide area and include municipalities covering rural areas. We will use population density to select neighbourhood to use in the model

Figure 2: Map of Greater Hobart showing all neighbourhoods



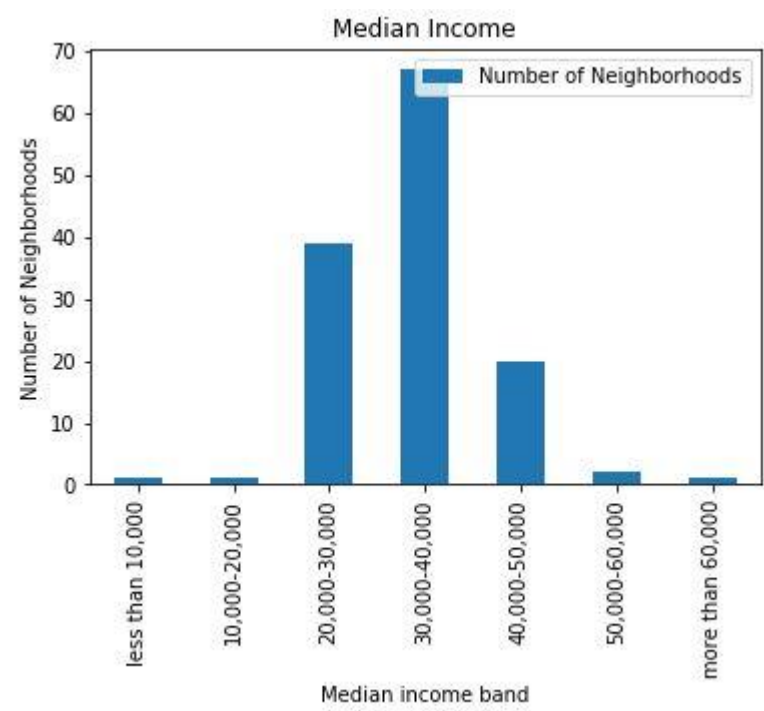
4.2. Analysis and visualisation by median income

Figure 3: Basic descriptive statistics of median income

	median_income
count	131.000000
mean	33614.625954
std	8483.050072
min	0.000000
25%	28548.000000
50%	33644.000000
75%	38480.000000
max	69836.000000

The Median income for the region is AUD33,644 and Mean income for the region is AUD33,614. The highest neighbourhood median income for the region is AUD69,836 while the lowest was 0.0. Median income band \$30,000 – 40,000 has most neighbourhoods (figure 4). There were 66 neighbourhoods above AUD33,614.62 mean income.

Figure 4: Number of neighbourhoods in each income band



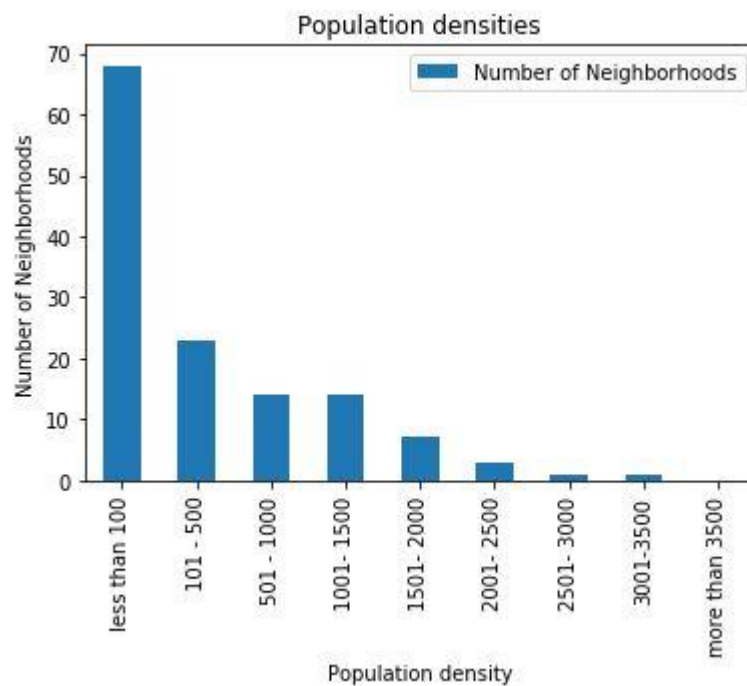
4.3 Analysis and visualisation by population density

Figure 5: Basic descriptive statistics of neighbourhood population

	population	area_sq_km	pop_density
count	131.000000	131.000000	131.000000
mean	1671.038168	17.369084	450.771286
std	2289.692859	29.542570	677.714237
min	0.000000	0.165000	0.000000
25%	225.500000	2.345000	17.928629
50%	657.000000	7.332000	87.174620
75%	2324.500000	17.033000	629.528486
max	11927.000000	208.351000	3357.575758

Total population for the region is 218906 with median population density of 87 persons/sq_km and mean population density of 450.77 persons/sq_km. The highest Population density for the region is 3,358 and the lowest is 0. There are 41 neighbourhoods above 451 mean population density

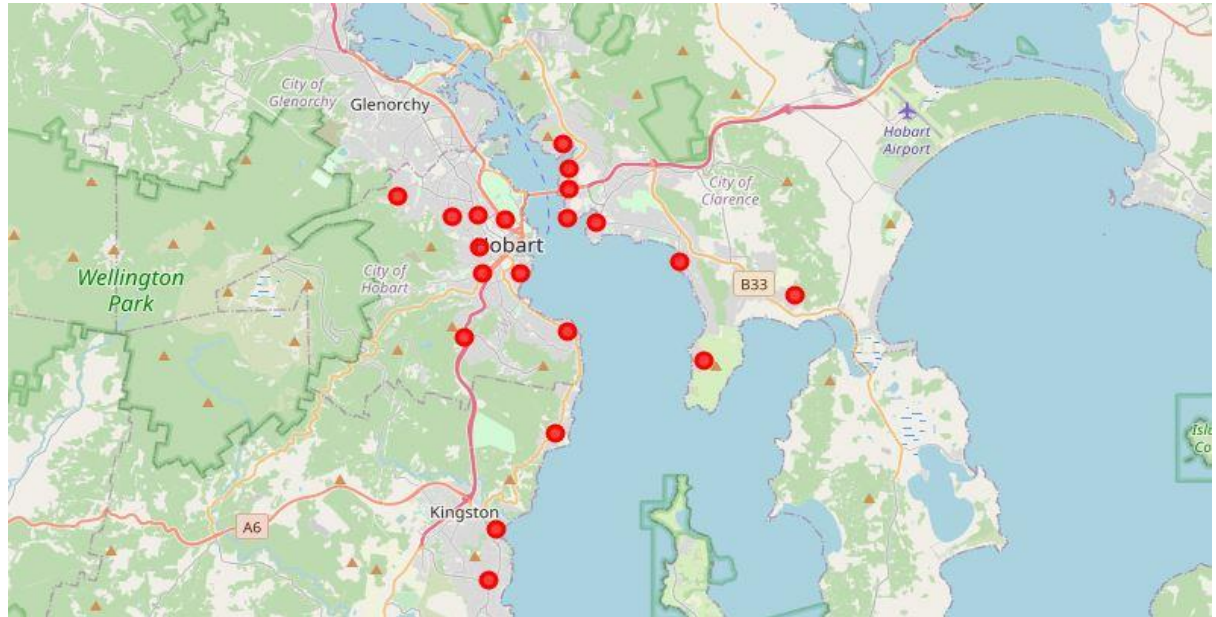
Figure 6: Number of neighbourhoods in population density band



4.4 Selecting neighbourhoods for modelling

There were 20 neighbourhoods with population density above 450.77 persons/sq_km **and** median income above AUD33,614.62. These were chosen for further analysis and clustering. As shown on the map (figure 6) these neighbourhoods are close to the city of Hobart.

Figure 6: map of high income and densely populated neighbourhoods



4.5 Venues data from Foursquare

782 venues were retrieved from the selected 20 neighbourhoods. Neighbourhoods tend to cover a wide area hence a radius of 1km was used. Results were limited to 100. There are 87 unique categories. As shown on figure 8, some neighbourhoods e.g. Tranmere, Taroona, have very few venues. Figure 8 shows dataframe head(10) of neighbourhoods and top 10 venues.

Figure 7: Dataframe head of neighbourhoods and top 10 most common venues

Neighborhood	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 Battery Point	Café	Coffee Shop	Hotel	Park	Italian Restaurant	Indian Restaurant	Bakery	Bar	Pub	Gastropub
1 Bellerive	Café	Coffee Shop	Hotel	Park	Italian Restaurant	Indian Restaurant	Bakery	Bar	Pub	Gastropub
2 Blackmans Bay	Café	Grocery Store	Australian Restaurant	Bar	Coffee Shop	Beach	Fish & Chips Shop	Yoga Studio	Food Truck	Dessert Shop
3 Glebe	Café	Park	Pub	Coffee Shop	Indian Restaurant	Hotel	Italian Restaurant	Australian Restaurant	Bakery	Bar
4 Howrah	Grocery Store	Liquor Store	Home Service	Athletics & Sports	Shopping Mall	Gym	Pharmacy	Field	Bus Station	Gastropub
5 Kingston Beach	Grocery Store	Fast Food Restaurant	Pizza Place	Supermarket	Fish & Chips Shop	Café	Shopping Mall	Beach	Coffee Shop	Bar
6 Lenah Valley	Bakery	Asian Restaurant	Café	Food Truck	Department Store	Dessert Shop	Distillery	Farmers Market	Fast Food Restaurant	Field
7 Lindisfarne	Park	Harbor / Marina	Pub	Trail	Grocery Store	Gas Station	Bakery	Supermarket	Bay	Fast Food Restaurant
8 Montagu Bay	Harbor / Marina	Café	Department Store	Grocery Store	Fast Food Restaurant	Bakery	Pub	Coffee Shop	Dessert Shop	Juice Bar
9 Mount Stuart	Café	Bakery	Japanese Restaurant	Hotel	Pizza Place	Asian Restaurant	Dessert Shop	Italian Restaurant	Indian Restaurant	Park

Figure 8: Data frame of all the neighbourhoods and venues details

31:

	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Neighborhood						
Battery Point	81	81	81	81	81	81
Bellerive	81	81	81	81	81	81
Blackmans Bay	8	8	8	8	8	8
Glebe	94	94	94	94	94	94
Howrah	18	18	18	18	18	18
Kingston Beach	16	16	16	16	16	16
Lenah Valley	4	4	4	4	4	4
Lindisfarne	13	13	13	13	13	13
Montagu Bay	36	36	36	36	36	36
Mount Stuart	72	72	72	72	72	72
North Hobart	78	78	78	78	78	78
Oakdowns	2	2	2	2	2	2
Rose Bay	29	29	29	29	29	29
Rosny	41	41	41	41	41	41
Sandy Bay	20	20	20	20	20	20
South Hobart	79	79	79	79	79	79
Taroona	4	4	4	4	4	4
Tolmans Hill	4	4	4	4	4	4
Tranmere	2	2	2	2	2	2
West Hobart	100	100	100	100	100	100

4.6 KMeans Clustering

Venues columns were converted into dummy values by one hot encoding. These were then grouped and mean calculated, producing a dataframe (Figure9). This dataframe was used for KMeans clustering using 5 number of clusters resulting in result in figure 10. A map of neighbourhoods with cluster labelled shown.

Figure 9: Neighbourhood's grouped venues dataframe

121:

	Neighborhood	Asian Restaurant	Assisted Living	Athletics & Sports	Australian Restaurant	Bagel Shop	Bakery	Bar	Bay	Beach	Bistro	Bookstore	Breakfast Spot	Brewery	Burger Joint	Bus Station	Business Service	Café	Chinese Restaurant	Cocktail Bar	Coffee Shop	Construction & Landscaping	Convenience Store	Ci Gr
0	Battery Point	0.012195	0.0	0.000000	0.012195	0.012195	0.036585	0.036585	0.0	0.000000	0.012195	0.012195	0.024390	0.0	0.024390	0.000000	0.0	0.097561	0.0	0.024390	0.085386	0.0	0.0	0.0
1	Bellerive	0.032258	0.0	0.000000	0.032258	0.000000	0.032258	0.000000	0.0	0.032258	0.000000	0.000000	0.000000	0.0	0.000000	0.000000	0.0	0.064516	0.0	0.000000	0.000000	0.0	0.0	0.0
2	Blackmans Bay	0.000000	0.0	0.000000	0.125000	0.000000	0.000000	0.125000	0.0	0.125000	0.000000	0.000000	0.000000	0.0	0.000000	0.000000	0.0	0.250000	0.0	0.000000	0.125000	0.0	0.0	0.0
3	Glebe	0.010638	0.0	0.000000	0.042553	0.010638	0.042553	0.031915	0.0	0.000000	0.010638	0.021277	0.010638	0.0	0.021277	0.000000	0.0	0.117021	0.0	0.010638	0.053191	0.0	0.0	0.0
4	Howrah	0.000000	0.0	0.055556	0.000000	0.000000	0.055556	0.000000	0.0	0.055556	0.000000	0.055556	0.000000	0.0	0.000000	0.055556	0.0	0.000000	0.0	0.000000	0.000000	0.0	0.0	0.0

venues_grouped = use_join(venues_grouped.set_index('Neighborhood'), on='Neighborhood')

Figure 10: Clustering algorithm and array of clusters

KMeans Clustering

```
# set number of clusters
kclusters = 5

hobart_grouped_clustering = venues_grouped.drop('Neighborhood', 1)

# run k-means clustering
kmeans = KMeans(n_clusters=kclusters, random_state=0).fit(hobart_grouped_clustering)

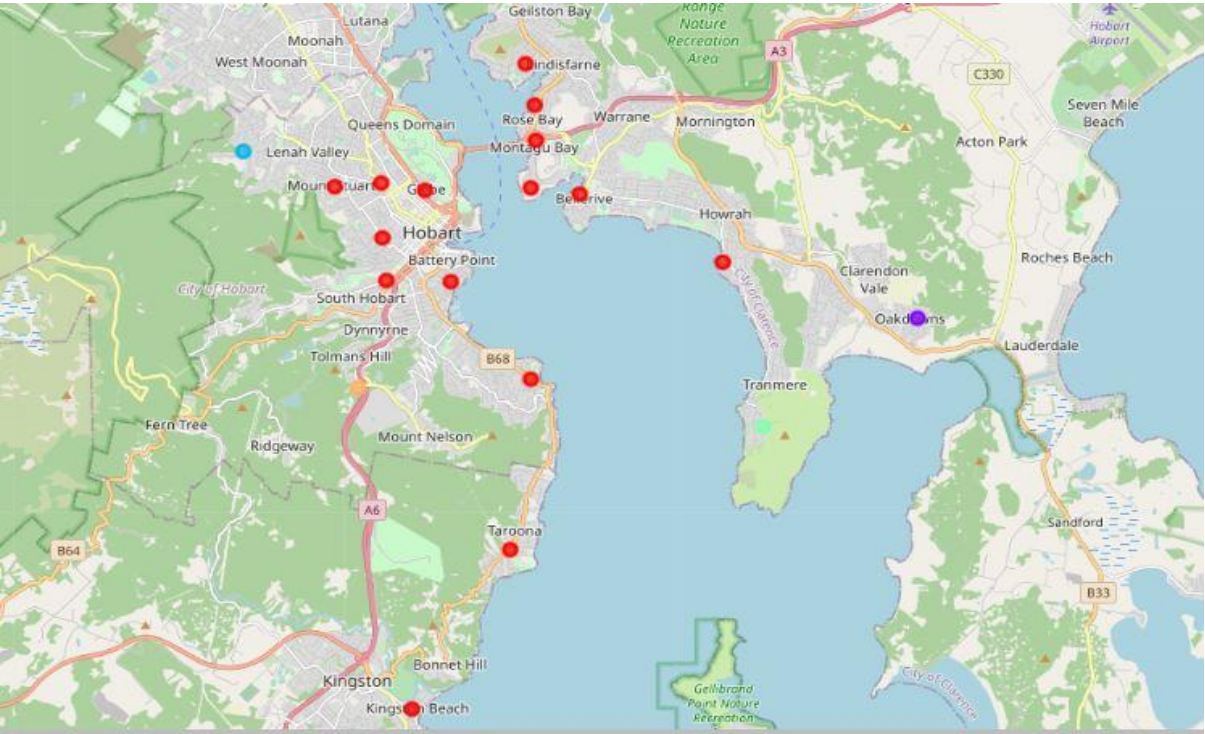
# check cluster Labels generated for each row in the dataframe
kmeans.labels_[0:]

array([0, 0, 0, 0, 0, 0, 2, 0, 0, 0, 0, 1, 0, 0, 0, 0, 4, 3, 0])
```

Figure 11: Dataframe neighbourhood, 10 most common venues and cluster label

	Neighborhood	postcode	LGA	latitude	longitude	median_income	pop_density	Cluster_Labels	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	Battery Point	7004	Hobart	-42.89277	147.33324	47164.0	2949.778434	0	Café	Coffee Shop	Hotel	Park	Italian Restaurant	Indian Restaurant	Bakery	Bar	Pub	Gastropub
1	Bellerive	7018	Clarence	-42.87482	147.36763	35308.0	1593.142453	0	Café	Coffee Shop	Hotel	Park	Italian Restaurant	Indian Restaurant	Bakery	Bar	Pub	Gastropub
2	Blackmans Bay	7052	Kingborough	-42.99887	147.31895	36816.0	1151.490733	0	Café	Grocery Store	Australian Restaurant	Bar	Coffee Shop	Beach	Fish & Chips Shop	Yoga Studio	Food Truck	Dessert Shop
3	Glebe	7000	Hobart	-42.87403	147.32635	43732.0	3357.575758	0	Café	Park	Pub	Coffee Shop	Indian Restaurant	Hotel	Italian Restaurant	Australian Restaurant	Bakery	Bar
4	Howrah	7018	Clarence	-42.88869	147.40568	37440.0	1323.887873	0	Grocery Store	Liquor Store	Home Service	Athletics & Sports	Shopping Mall	Gym	Pharmacy	Field	Bus Station	Gastropub

Figure 12: Map of neighbourhoods showing cluster label



4.7 Examining clusters

Clusters were examined. Cluster 1 was identified as café/coffee shop group with good variety of venues. However, there are neighbourhoods in the cluster where café is not the most common venue. These are shown in figure 15. Other clusters had only one member. This may indicate that the chosen neighbourhoods are similar.

Figure 13: Cluster 1 Dataframe

Cluster 1

```
3]: hobart_merged.loc[hobart_merged['Cluster_Labels'] == 0]
```

```
3]:
```

	Neighborhood	postcode	LGA	latitude	longitude	median_income	pop_density	Cluster_Labels	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	Battery Point	7004	Hobart	-42.89277	147.33324	47164.0	2949.778434	0	Café	Coffee Shop	Hotel	Park	Italian Restaurant	Indian Restaurant	Bakery	Bar	Pub	Gastropub
1	Bellerive	7018	Clarence	-42.87482	147.36763	35308.0	1593.142453	0	Café	Coffee Shop	Hotel	Park	Italian Restaurant	Indian Restaurant	Bakery	Bar	Pub	Gastropub
2	Bladsmans Bay	7052	Kingborough	-42.99887	147.31895	36816.0	1151.490733	0	Café	Grocery Store	Australian Restaurant	Bar	Coffee Shop	Beach	Fish & Chips Shop	Yoga Studio	Food Truck	Dessert Shop
3	Glebe	7000	Hobart	-42.87403	147.32635	43732.0	3357.575758	0	Café	Park	Pub	Coffee Shop	Indian Restaurant	Hotel	Italian Restaurant	Australian Restaurant	Bakery	Bar
4	Howrah	7018	Clarence	-42.88969	147.40568	37440.0	1323.887873	0	Grocery Store	Liquor Store	Home Service	Athletics & Sports	Shopping Mall	Gym	Pharmacy	Field	Bus Station	Gastropub
5	Kingston Beach	7050	Kingborough	-42.98101	147.32286	34424.0	1340.067340	0	Grocery Store	Fast Food Restaurant	Pizza Place	Supermarket	Fish & Chips Shop	Café	Shopping Mall	Beach	Coffee Shop	Bar
7	Lindisfarne	7015	Clarence	-42.84772	147.35295	35568.0	889.704824	0	Park	Harbor / Marina	Pub	Trail	Grocery Store	Gas Station	Bakery	Supermarket	Bay	Fast Food Restaurant
8	Montagu Bay	7018	Clarence	-42.86351	147.35581	33644.0	1466.517857	0	Harbor / Marina	Café	Department Store	Grocery Store	Fast Food Restaurant	Bakery	Pub	Coffee Shop	Dessert Shop	Juice Bar
9	Mount Stuart	7000	Hobart	-42.87312	147.30230	44616.0	2320.116054	0	Café	Bakery	Japanese Restaurant	Hotel	Pizza Place	Asian Restaurant	Dessert Shop	Italian Restaurant	Indian Restaurant	Park
10	North Hobart	7000	Hobart	-42.87232	147.31458	37544.0	2449.901768	0	Café	Coffee Shop	Pub	Hotel	Australian Restaurant	Bakery	Park	Grocery Store	Indian Restaurant	Dessert Shop

Figure 14: statistics showing most grouped most common venues

```
5]:
```

	Neighborhood	LGA	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
count	16	16	16	16	16	16	16	16	16	16	16	16
unique	16	3	8	9	10	11	15	13	13	15	13	11
top	Bellerive	Hobart	Café	Coffee Shop	Pub	Park	Italian Restaurant	Indian Restaurant	Bakery	Bar	Indian Restaurant	Gastropub
freq	1	7	8	5	5	3	2	2	3	2	2	3

Figure 15: Neighbourhoods where café/coffee shop is not on top 2 most common venue

```
: 4      Howrah
5      Kingston Beach
7      Lindisfarne
8      Montagu Bay
12     Rose Bay
13     Rosny
14     Sandy Bay
16     Taroonna
Name: Neighborhood, dtype: object
```

5. RESULTS

Our model shows that cluster 2 is cafe and coffee shop region with a good variety of venues. Unsurprisingly these neighbourhoods are close to the city centre. Below are recommended neighbourhoods from this cluster where cafe/coffee shop is not in the top 2. These neighbourhoods have high population density, above median incomes and similar in terms of venues. These neighbourhoods are Howrah, Kingston Beach, Lindisfarne, Montagu Bay, Rose Bay, Rosny, Sandy Bay and Taroona.

6. DISCUSSION AND CONCLUSION

Using neighbourhood income, population density and venues data we recommended some neighbourhoods to open a cafe/coffee shop. From the result it shows neighbourhoods in inner Hobart are homogenous. We hope this helps investors shortlist locations better. The model can be improved further for example by using foot traffic in the analysis and using businesses/workplaces that do not appear in the Foursquare data.

7. REFERENCES

1. AustralianTownList.com - <https://www.australiantownlist.com/>
2. Australian Bureau of Statistics - <https://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/6524.0.55.0022011-12%20to%202016-17?OpenDocument>