Viability of Locations for Opening A Restaurant

28th August 2020

1 Introduction

1.1. Background

The metropolitan area of Greater Melbourne is divided into smaller areas called cities, which are subdivided into suburbs. A city is essentially a collection of neighbouring suburbs.

The population of Greater Melbourne is diverse in background and culture. When the area is examined at finer granularity, this diversity dissipates somewhat and clusters of similar cultures can be seen at the suburb- or sometimes city-level.

These clusters are often tied to the demand for certain types of food in the area. For example, you may encounter more Southeast Asian Restaurants in a city with a relatively large proportion of Southeast Asian residents.

1.2. Problem

The aim of this project is to determine the best areas in which to open a restaurant, and the type of restaurant that may experience more success than others in these areas.

2 Data Acquisition

2.1. Data Sources and Feature Selection

This project uses a list of Melbourne cities and suburbs from Wikipedia (https://en.wikipedia.org/wiki/List_of_Melbourne_suburbs)
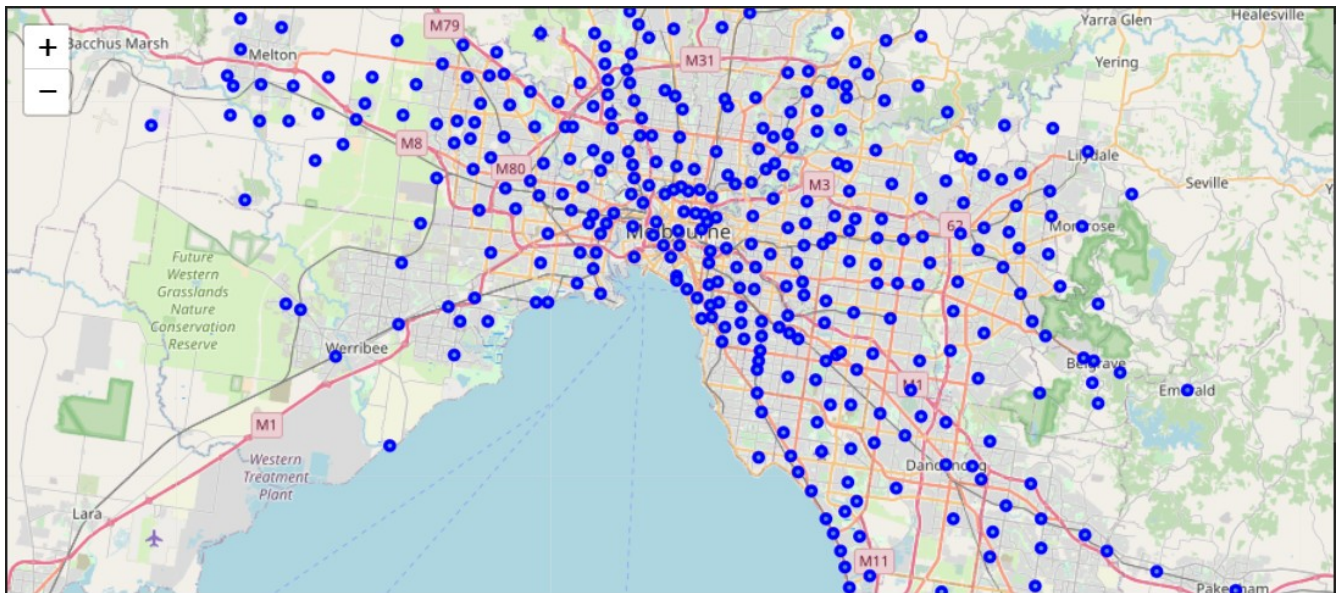
Figure 1: Greater Melbourne Suburbs

To understand demand for certain cuisines in different suburbs, the project looks at restaurant data from FourSquare. Specifically, we look at sub-categories of the top-level category "Food", while ignoring further sub categorization. For example, the category "Asian Restaurant" falls under "Food", and has many sub-categories such as "Chinese Restaurant" and "Vietnamese Restaurant"; we only observe the top-level "Asian Restaurant" as finer granularity may obfuscate the trends we are hope to find in our analysis.

This project utilizes population age and income data from the Australian Bureau of Statistics' (ABS) census. Age data is provided at the granularity of years (0yr, 1yr, 2yr… 99yr, above 99yr), a sample of which can be seen in Figure 2. This project converts these into age-groups (0-15yr, 16-27yr, 28-55, and 56 and older), a corresponding sample of which can be seen in Figure 3. These groups represent the different general types of customers. For example, the '56 and older' group represents retirees.

SSC_CODE_2016	0	1	2	3	4	4	5	6	7	8	...	76	77	78	79	79	84	89	94	99	over
SSC20001	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
SSC20002	86	82	44	57	44	319	33	40	34	33	...	17	21	12	15	93	91	62	28	14	0
SSC20003	34	30	41	30	50	179	40	47	59	55	...	9	17	4	18	74	57	57	26	8	0
SSC20004	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
SSC20005	0	0	0	0	0	0	0	0	0	3	...	0	0	3	4	6	5	4	0	0	0

Figure 2: Raw ABS Age Data

SSC_CODE_2016	0-15	16-27	28-55	56 and older
SSC20001	0	0	0	0
SSC20002	1310	2883	9591	3682
SSC20003	1521	1330	3053	2745
SSC20004	0	0	0	0
SSC20005	52	20	70	259

Figure 3: Grouped ABS Age Data

As an estimation of disposable income, this project uses the difference between income and mortgage repayments (Figure 5), both taken from the ABS census data (Figure 4).

	Median_mortgage_repay_monthly	Median_tot_hhd_inc_weekly	SSC_NAME_2016
20001	0	0	Abbeyard
20002	2142	2001	Abbotsford
20003	2200	2200	Aberfeldie
20004	0	0	Aberfeldy
20005	1424	1109	Acheron

Figure 4: ABS Income and Mortgage Data

	SSC_NAME_2016	Disposable income monthly
20002	Abbotsford	6552.84525
20003	Aberfeldie	7359.55000
20005	Acheron	3394.88225
20007	Adams Estate	2953.64225
20008	Addington	1881.08050

Figure 5: Disposable income as function of income and mortgage repayments.

2.2. Caveats

We obtain FourSquare data based on a 500m radius of suburb geolocation centers. This works for most suburbs as it tends to include the suburb's commercial center. However, there may be cases of suburbs with commercial centers closer to its borders, and outside of a 500m radius of it's geological center. For the purposes of this project, and as we must adhere to cost-free limits on data from FourSquare, we will not account for this case.

Another caveat on FourSquare data is its limit on results (2 pages of 50 results each) for cost-free usage. This seems to be an issue only for suburbs in the Central Business District, where there are more than 100 Food venues in a 500m radius.

3   Data Analysis

In order to gauge the suitability of locations for opening a restaurant, this project looks at several factors including:

- Mean disposable income in the area
- Frequency of various restaurant cuisines in the area
- Age distribution in the area

Mean Disposable Income

To narrow down the list of suburbs to look at, we first rank all suburbs by disposable income descending. We then find out which cities the top 50 suburbs belong to, which gives us a list of 20 cities. The list is then culled to the top 10 for further analysis. Figure 6 (below) illustrates which cities the 50 suburbs belong to, along with the mean disposable income of these suburbs.
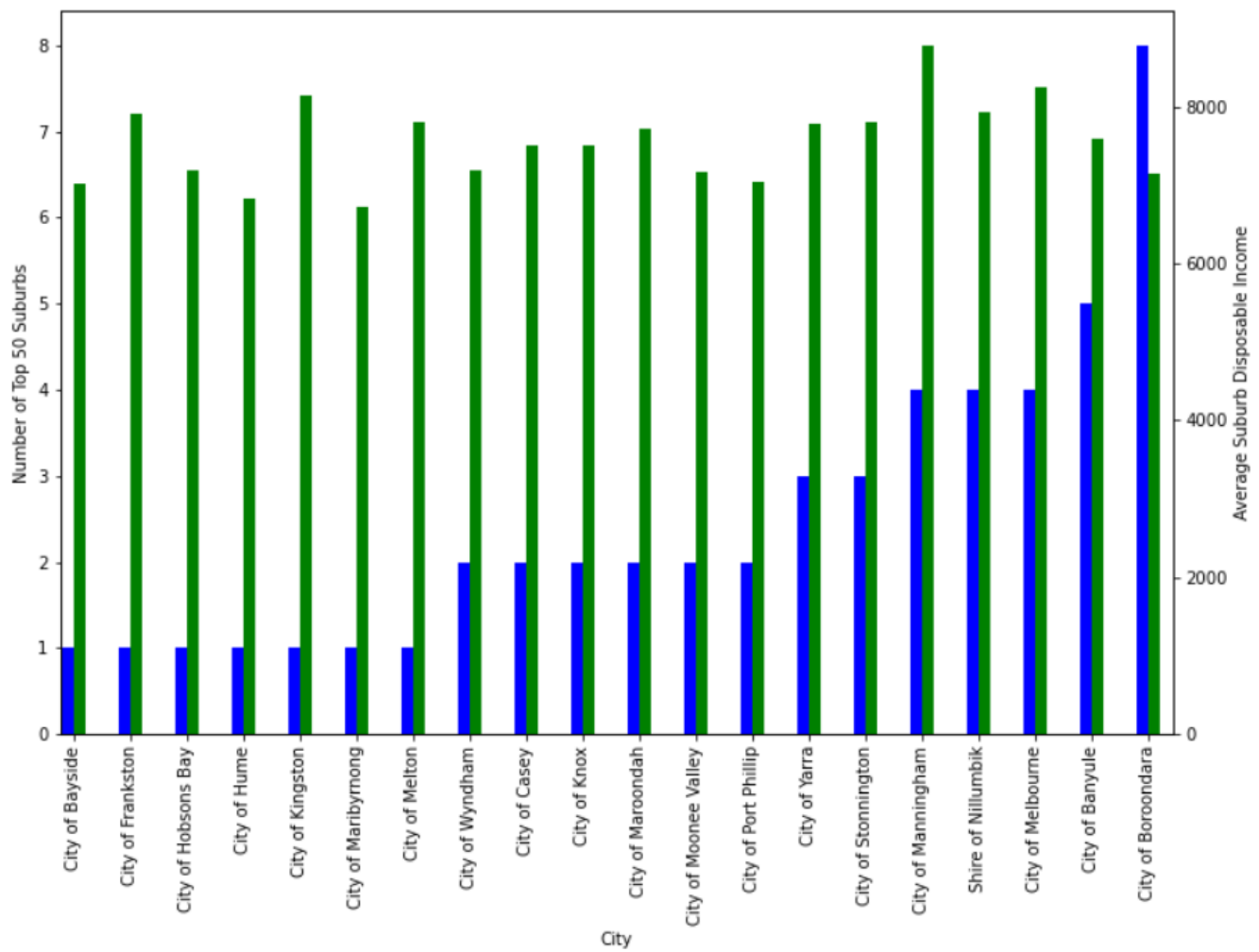
Figure 6: Mean Disposable Income of Suburbs, by City

The next factor we look at is the popularity of restaurant cuisines. Figure 7 shows a word map to give a rough idea of popularity for the entire greater Melbourne region.

The following stacked bar chart (Fig. 8) gives an overview of the top 3 cuisines in each of the top 10 cities.

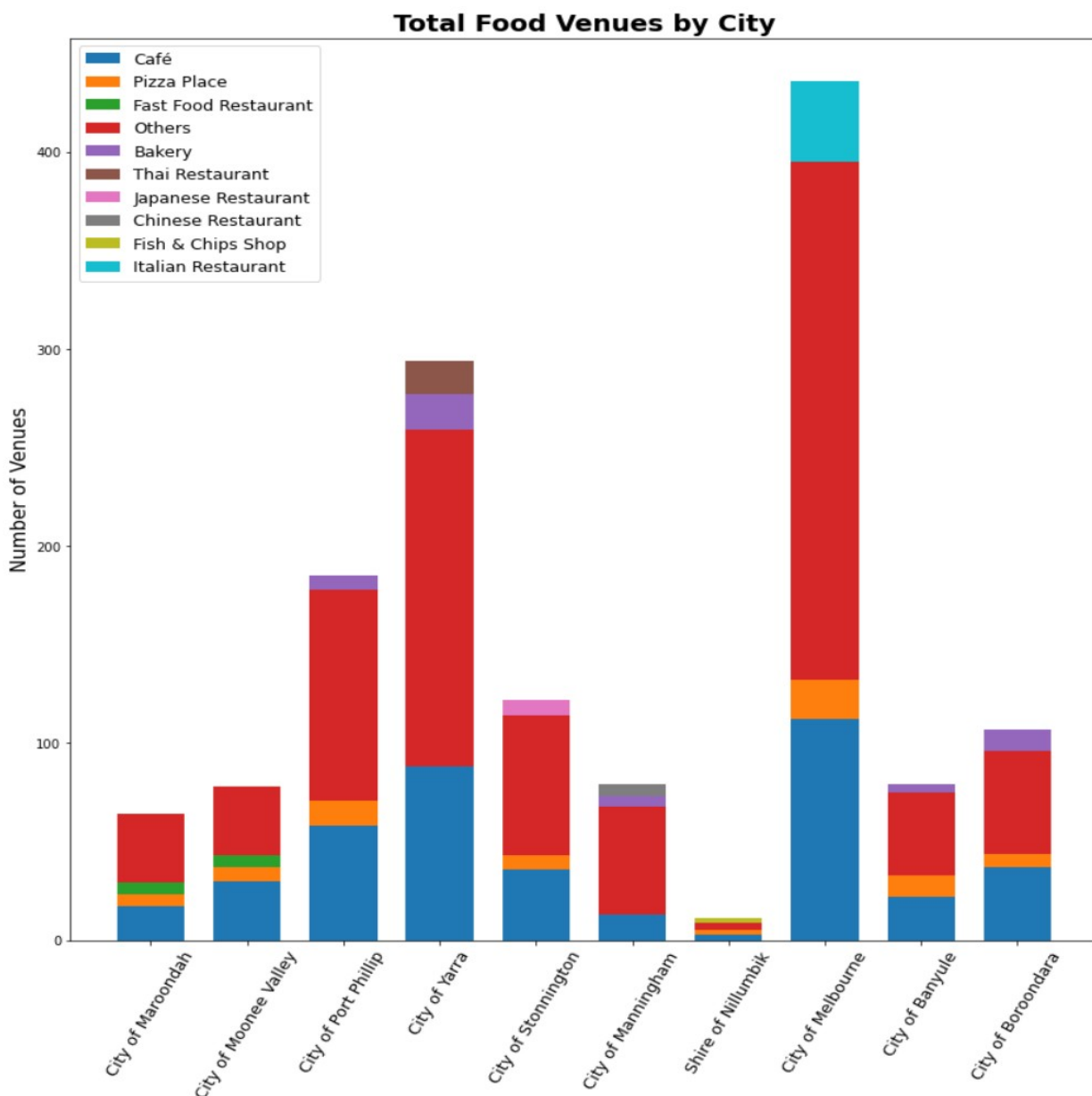*Figure 7: Cuisine Commonality Word Map*

Figure 8: Cuisine frequencies, by city

The above bar chart tells us that Cafes and Pizzarias are pervasive. Opening a restaurant of a cuisine that is common is somewhat of a double-edged sword, in that while demand for that cuisine will almost certainly exist, competition may be stiff and the restaurant must have some other means of standing out.

To understand the distinct demands for different cuisines in each city, we remove Cafes and Pizzarias from the bar charts (Fig. 9):
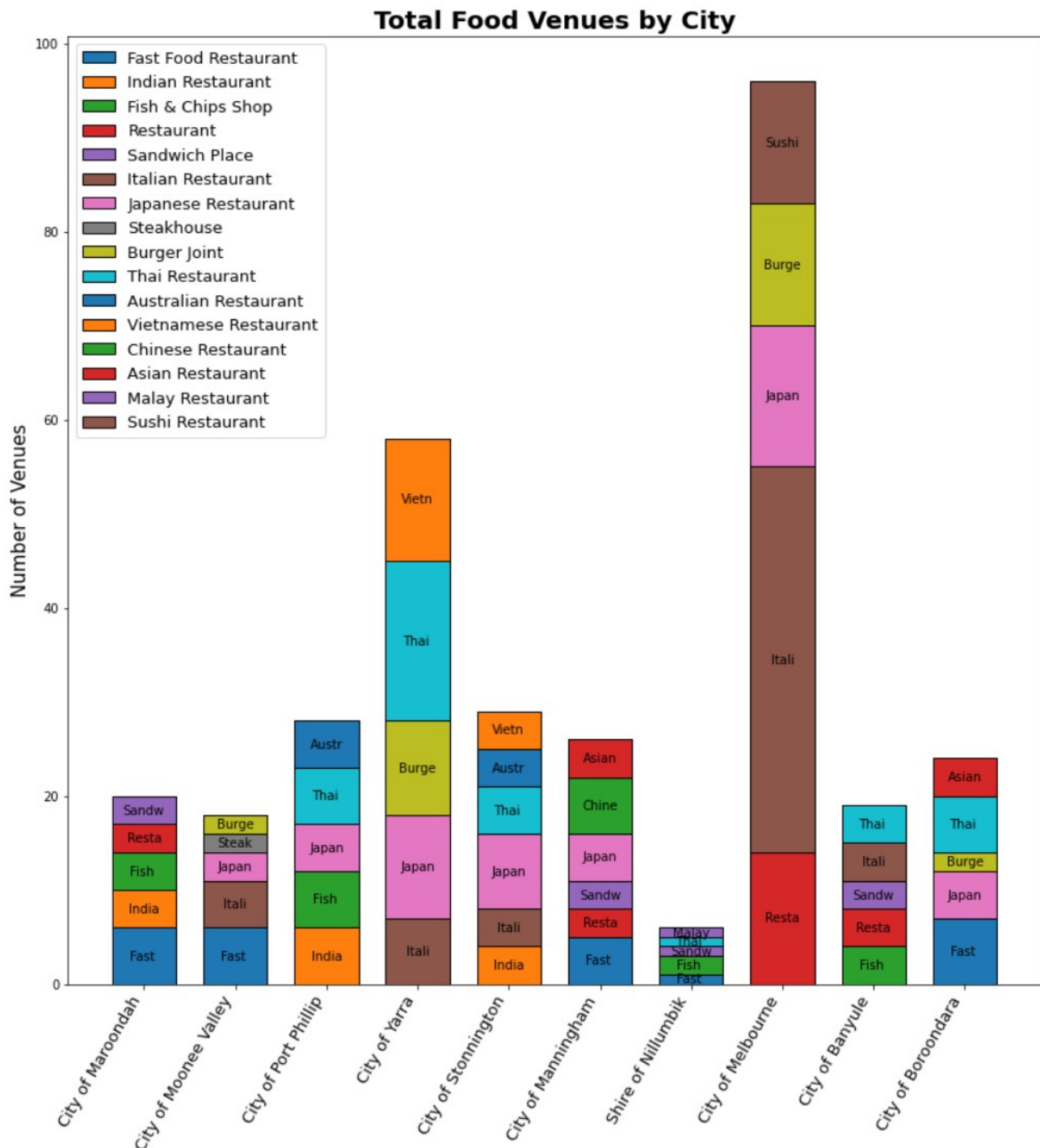

Figure 9: Finer Detail of Cuisine Popularity, by City

Knowing the common types of restaurants in a particular area gives us a sense of what demands in that area are. For example, in the City of Port Phillip – Fish & Chips, Indian, Thai and Japanese food have their established fans.

Opening a restaurant serving up one of these four cuisines might at best be a safe investment with less than spectacular returns, and at worst will leave the restaurant struggling to differentiate itself from the incumbents. However, knowing the cuisines gives the business starter some idea of the tastes of potential customers in the area.

Population by Age

The next factor we look at is the age distribution of each top-10 city. Figures 10 and 11 give us an overview.
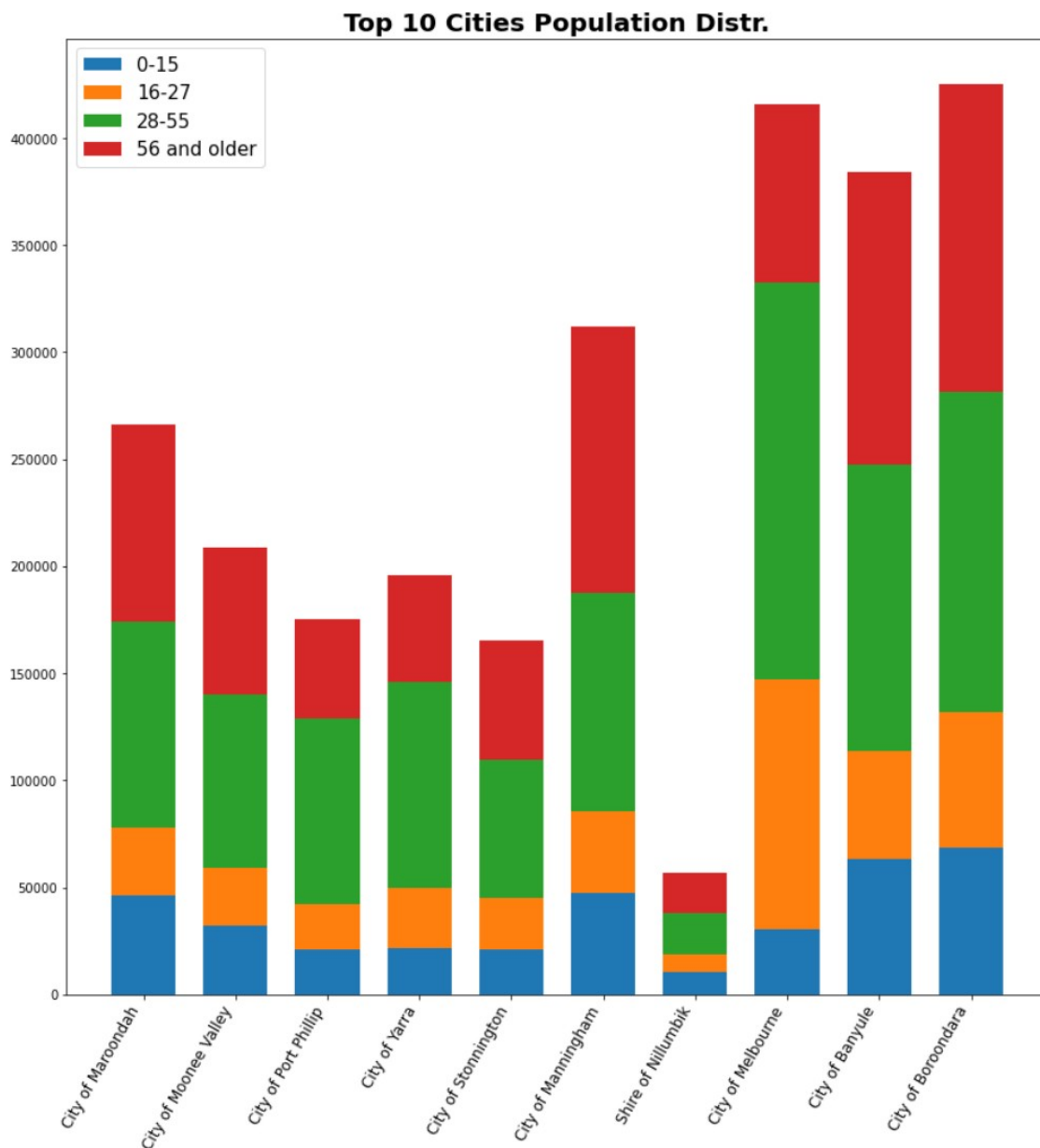

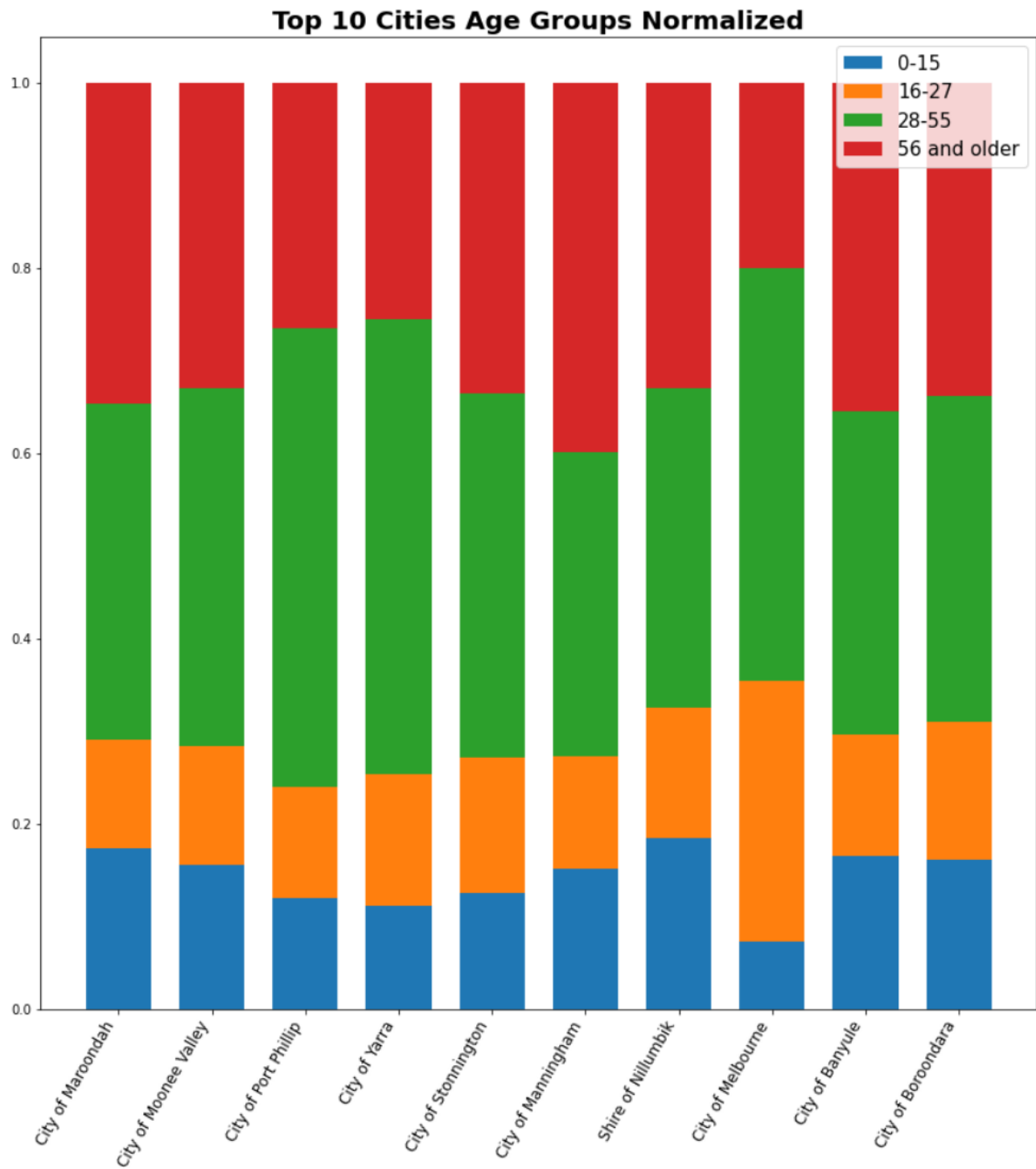*Figure 10: Age distribution, by city*

Figure 11: Age distribution, by city, normalized

Some observations from the age distribution charts (Figs. 10 & 11):

- *Shire of Nillumbik* has a relatively low population. This is likely the reason it has relatively few restaurants, as observed in figures 8 and 9.

- Cities of Melbourne, Banyule, and Boroondara have distinctly larger proportions of 16-27 year-olds compared to the other cities. This is indicative of families with grown children. A restaurant in these areas might find a competitive edge in catering to that generation of youths.

- Cities of Maroondah, Manningham, and again Banyule and Boroondah have larger proportions of children 0-15 years of age. This likely means there are a relatively high number of young families in the area. A restaurant in these areas can leverage that by creating a family-friendly ambiance to attract these families of diners. Interestingly, these four cities also seem to have the highest proportions of elderly residents.

Exploring Alternative Cities

Once a city and restaurant type has been chosen, it may be worthwhile looking at alternative cities with similar characteristics. Tables 1 and 2 are the results of applying KMeans on both age groups and restaurant types to cluster the cities.

Table 1: City Clusters by Age Groups

Cluster 0	Cluster 1	Cluster 2
- Hume - Bayside - Casey - Cardinia - Whittlesea	- Boroondara - Brimbank - Monash - Glen Eira - Greater Dandenong - Hobsons Bay - Moonee Valley - Wyndham - Melton - Whitehorse	- Melbourne - Stonnington - Manningham - Maribyrnong
Cluster 3	**Cluster 4**	
- Mornington Peninsula - Banyule - Maroondah - Nillumbik - Knox - Kingston - Frankston - Yarra Ranges	- Moreland - Yarra - Darebin - Port Phillip	

Table 2: City Clusters by Restaurant Types

Cluster 0	Cluster 1	Cluster 2
• Banyule • Bayside • Brimbank • Casey • Darebin • Frankston • Greater Dandenong • Hobsons Bay • Hume • Knox • Manningham • Maroondah • Melton • Monash • Whitehorse • Whittlesea • Wyndham • Cardinia • Mornington Peninsula • Nillumbik • Yarra Ranges	• Melbourne • Yarra	• Boroondara • Glen Eira • Kingston • Maribyrnong • Moonee Valley • Moreland • Port Phillip • Stonnington

4   Conclusion

This study was intended to find an ideal location and respective type of restaurant to open. Upon analysis however, it appears the problem definition must be narrowed for a more definitive answer to exist. While a single correct answer does not exist for the initial question, this study has led to a greater understanding of market demand differences in different areas within the greater Melbourne metropolitan region. It may therefore arguably be a starting point for anyone asking the initial question.