

The Classification of Suburbs in Victoria, Australia

Angel Wu

18/01/20

1. Introduction

In this project we will try to compare different suburbs in Victoria, Australia. Specifically, this report will be targeted to people interested in moving into one of the suburbs in Victoria, Australia.

Since there are lots of suburbs in Victoria and the purpose of this project is to compare the different types of suburbs only, we will group the suburbs based on the occurrences of venues from different categories, this will then be converted into a easily understandable visual representation that shows the distribution of different types of suburbs so that individuals interested in this project can choose the best type of suburb that suits their needs.

2. Data

Following data sources will be needed to extract/generate the required information:

- The names and geographical locations of the suburbs in Victoria, Australia will be obtained using a csv file called 'Australian_Post_Codes_Lat_Lon.csv'. The csv file could be retrieved from the following link: <http://www.corra.com.au/australian-postcode-location-data/>
- Coordinate of the central city of the Victoria State, Melbourne will be obtained using Google Maps API geocoding.
- The venues in each suburb, their associated categories and their geographical location will be obtained using Foursquare API.

Data cleaning

Since we are only interested in the suburbs in Victoria, Australia, we should reduce the number of rows in the dataframe by only selecting the entries that has the value 'VIC' (which represents the Victoria State) under the column 'state'. To prevent the overlapping data points in our final data visualisation step, any rows with duplicating values for the 'suburb' column is dropped. Since only the suburb name and the geographical location of the suburbs are needed, we only select the columns 'suburb', 'lat' and 'lon' for the completion of this project. Since the original dataset is quite large, for demonstration purpose, only the first 50 rows, that is, the first 50 suburbs are chosen for analysis.

Methodology

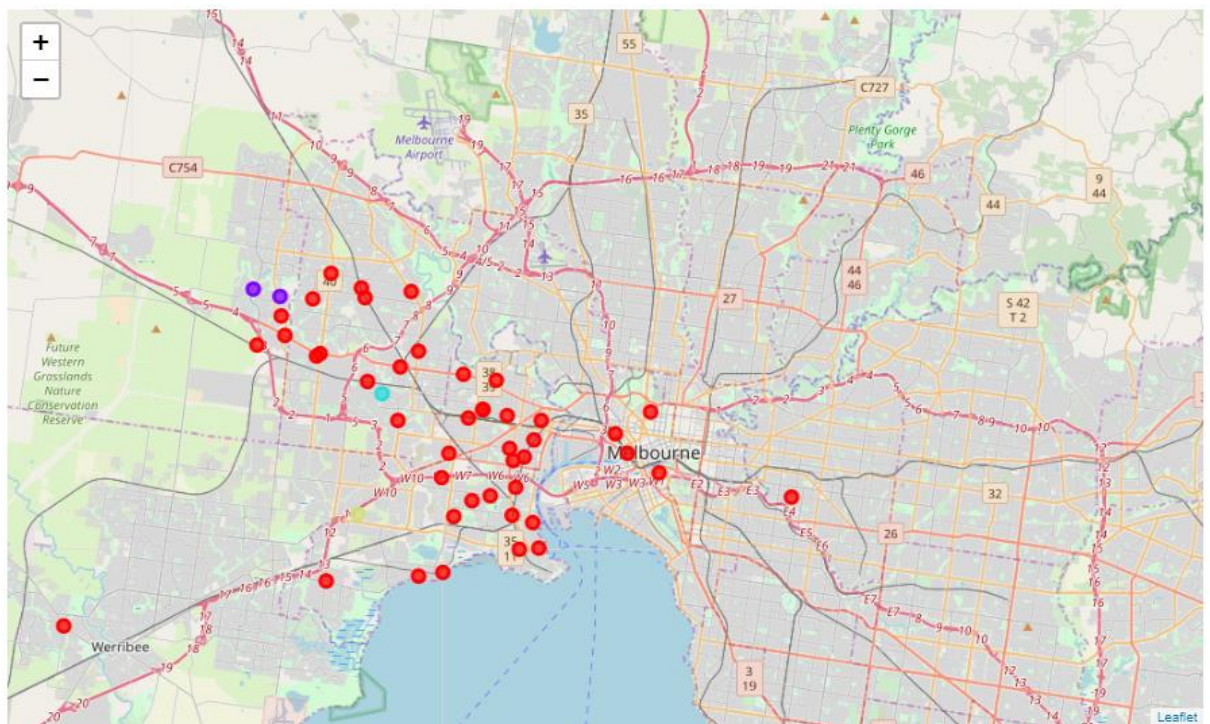
In first step we have collected the required data: the geographical locations and the categories of every venues within 1km from the centre of each suburb by utilising the Foursquare API.

Second step in our analysis will be creating clusters of suburbs that have certain level of commonality, this is determined by calculating the mean of the frequency of occurrence of venues in each suburb.

The third and final step will be presenting the clustering of suburbs by utilising the Map function from folium, so that the clients can make more informed decision about which suburb will suit their lifestyle.

Result and Discussion

The following is a screenshot representing the clustering of suburbs based on our calculation on the frequency of occurrence of venues for each suburb:



The clustering of suburbs

The following is a subset of the suburbs that is being categorised as cluster 0.

	suburb	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	WEST MELBOURNE	Café	Platform	Train Station	Theater	Food Court	Bus Stop	Bar	Wine Shop	Australian Restaurant	Bagel Shop
1	SOUTHBANK	Bar	Theater	Café	Italian Restaurant	Performing Arts Venue	Grocery Store	Art Gallery	Australian Restaurant	Steakhouse	Hotel
2	DOCKLANDS	Café	Coffee Shop	Restaurant	Hotel	Shopping Mall	Sandwich Place	Pizza Place	Pub	Bar	Japanese Restaurant
3	UNIVERSITY OF MELBOURNE	Café	Coffee Shop	Athletics & Sports	Pub	Hotel	College Cafeteria	Juice Bar	Lounge	Food Court	Electronics Store
4	FOOTSCRAY	Vietnamese Restaurant	Asian Restaurant	Café	Bakery	Platform	Bar	Coffee Shop	Sandwich Place	Chinese Restaurant	Light Rail Station
5	SEDDON	Café	Bakery	Wine Shop	Supermarket	Gastropub	Liquor Store	Park	Dance Studio	Pizza Place	Gym
6	SEDDON WEST	Grocery Store	Fish & Chips Shop	Gym	Playground	Department Store	Cupcake Shop	Thai Restaurant	Food & Drink Shop	Shopping Mall	Supermarket
7	BROOKLYN	Café	Electronics Store	Food Truck	Wine Shop	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Falafel Restaurant	Dumpling Restaurant	Food
8	KINGSVILLE	Café	Convenience Store	Miscellaneous Shop	Skate Park	Soccer Field	Fast Food Restaurant	Fish & Chips Shop	Sandwich Place	Supermarket	Thai Restaurant
9	KINGSVILLE WEST	Grocery Store	Fish & Chips Shop	Gym	Playground	Department Store	Cupcake Shop	Thai Restaurant	Food & Drink Shop	Shopping Mall	Supermarket
10	MAIDSTONE	Gym	Shopping Mall	Café	Latin American Restaurant	Electronics Store	Fast Food Restaurant	Farmers Market	Falafel Restaurant	Dumpling Restaurant	Flower Shop
11	TOTTENHAM	Accessories Store	Train Station	Liquor Store	Cupcake Shop	Department Store	Dessert Shop	Diner	Discount Store	Donut Shop	Flower Shop
12	WEST FOOTSCRAY	Indian Restaurant	Café	South Indian Restaurant	Grocery Store	Fried Chicken Joint	Electronics Store	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Falafel Restaurant
13	YARRAVILLE	Café	Pizza Place	Grocery Store	Fish & Chips Shop	Burger Joint	Cambodian Restaurant	Farmers Market	Lounge	Coffee Shop	Gift Shop
14	YARRAVILLE WEST	Sandwich Place	Skate Park	Miscellaneous Shop	Kids Store	Fish & Chips Shop	Farmers Market	Falafel Restaurant	Electronics Store	Dumpling Restaurant	Wine Shop
15	NEWPORT	Café	Pizza Place	Grocery Store	Ice Cream Shop	Convenience Store	Bagel Shop	Thai Restaurant	Park	Beer Garden	Dumpling Restaurant
16	SOUTH KINGSVILLE	Pizza Place	Fish & Chips Shop	Restaurant	Bakery	Dry Cleaner	Fast Food Restaurant	Farmers Market	Falafel Restaurant	Electronics Store	Dumpling Restaurant
17	SPOTSWOOD	Café	Bakery	Gas Station	Breakfast Spot	Train Station	Golf Course	Post Office	Pub	Dry Cleaner	Farmers Market

The common venues in cluster 0 show more variety in terms of the venues in vicinity. Apart from the different cuisines offered by the restaurants, there is also provision of public transports, such as train station and bus stops, this indicates that living in the suburbs in cluster 0 is suitable for those who prefer to take public transports. There are also more venues for entertainment, such as park, playground, dance studio, this suggests that these suburbs are suitable for families with children. The suburbs in cluster 0 are also quite convenient for living, since markets and shopping malls are considered as common venues.

Cluster 1:

	suburb	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
38	CAROLINE SPRINGS	Athletics & Sports	Wine Shop	Dumpling Restaurant	Flower Shop	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Falafel Restaurant	Electronics Store	Dry Cleaner
40	DEER PARK NORTH	Athletics & Sports	Child Care Service	Wine Shop	Dumpling Restaurant	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Falafel Restaurant	Electronics Store	Dry Cleaner

For the suburbs being categorised into cluster 1, as the most common venue belongs to the category of 'Athletics and sports', we may conclude that these suburbs are suitable for sporty individuals. The suburbs in cluster 1 are also convenient for living, since 'Farmers markets' are considered as common venues.

Cluster 2:

	suburb	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
26	GLENGALA	Pizza Place	Food	Deli / Bodega	Department Store	Dessert Shop	Diner	Discount Store	Donut Shop	Dry Cleaner	Dumpling Restaurant

Cluster 3:

	suburb	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
48	LAVERTON NORTH	Furniture / Home Store	Wine Shop	Dumpling Restaurant	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Falafel Restaurant	Electronics Store	Dry Cleaner	Food

Since there is only one suburb being categorised into cluster 2 and cluster 3, we can not identify a clear feature for suburbs in these clusters, in this case we will regard those suburbs as others.

Conclusion

The purpose of this project is to classify suburbs in Victoria based on the venues that are within 1km of each suburb centre. By showing the suburbs and labelling them according to the clusters on a map, the clients can have more comprehensive idea about their potential options and also the geographical locations of these suburbs. By showing the suburbs in each cluster and summarise the key features in each cluster in the Result section, the user can quickly grasp the key ideas in each suburb and choose their preferred suburbs according to their needs.