

Data Science Capstone - Project

Analysis of Neighborhoods -
Using data of Australia





Introduction/Business Problem

The idea is to analyse neighborhoods of Australia and try to cluster related neighborhoods together. Neighborhoods belonging to the same cluster can be analysed using their most common venues in order to take important decisions such as; Which neighborhood is better to open up a restaurant? Which neighborhoods prefer cafe's and cinemas or even nightclubs?

This will help any business to decide where they want to set up shop as they will know which neighborhood will prefer their business. It can also help to analyse if any particular type of business is not present in a given neighborhood like a neighborhood has only restaurants and cafes but no night clubs then somebody who owns a chain of nightclubs can set one up there.



Data Section

The data which will be used for the project is from the page https://en.wikipedia.org/wiki/Postcodes_in_Australia

It contains two tables;

1. A table that contains the state/territories along with their abbreviations
2. A table containing the postcodes along with locality and state/territory abbreviations

Both tables will be joined using the common column i.e. abbreviations so that it will be easier for the user to understand which state/territory the locality belongs to.

Final Dataset

	PostCode	State/Territory	Neighborhood	Latitude	Longitude
0	0872	Northern Territory	ERNABELLA, FREGON, INDULKANA, MIMILI, NGAANYAT...	-25.719898	131.957835
1	2406	New South Wales	MUNGINDI	-29.030752	149.188191
2	2540	New South Wales	HMAS CRESWELL, JERVIS BAY	-34.983059	150.603134
3	2611	Australian Capital Territory	COOLEMAN, BIMBERI, BRINDABELLA, URIARRA	-35.316175	149.010503
4	2620	New South Wales	HUME, KOWEN FOREST, OAKS ESTATE, THARWA, TOP NAAS	-35.277574	149.236242
5	3500	Victoria	PARINGI	-34.196090	142.142540
6	3585	Victoria	MURRAY DOWNS	-35.342298	143.558281
7	3586	Victoria	MALLAN	-35.402222	143.654334
8	3644	Victoria	BAROOGA, LALALTY	-35.911559	145.671953
9	3691	Victoria	LAKE HUME VILLAGE	-36.162044	146.961706
10	3707	Victoria	BRINGENBRONG	-36.176704	147.947410
11	4377	Queensland	MARYLAND	-28.561347	151.988089
12	4380	Queensland	MINGOOLA	-28.703805	151.911134
13	4825	Queensland	ALPURRURULAM	-20.723552	139.384069



Methodology Section

In this section, we used FourSquare API to get the nearest venues to each neighborhood of the dataset.

We then found the top 10 most common venues of each neighborhood then performed K means clustering to group similar neighborhoods together.

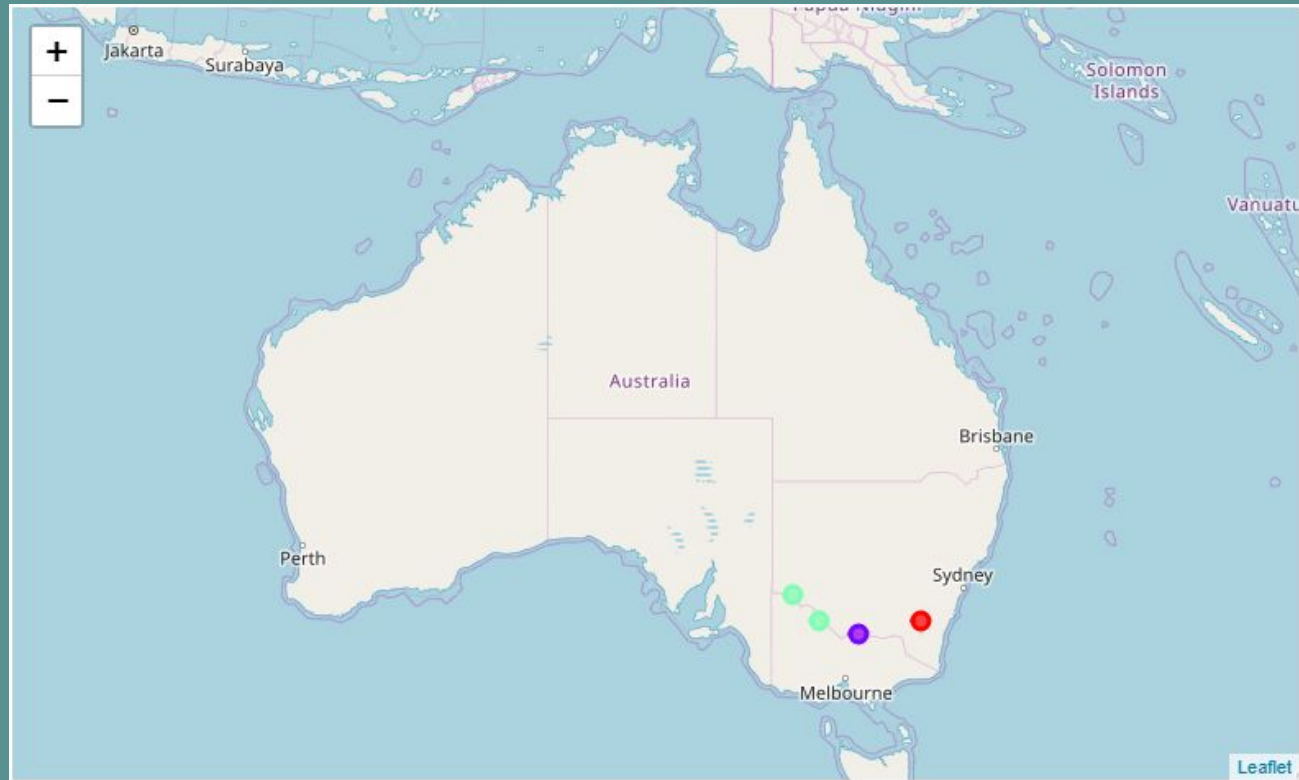
Dataset after processing...

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 Common Venue
BAROOGA, LALALTY	Café	Beach	Sandwich Place	Planetarium	Pizza Place	Italian Restaurant	Fish & Chips Shop	Fast Food Restaurant	Discount Store
COOLEMAN, BIMBERI, BRINDABELLA, URIARRA	Planetarium	Café	Sandwich Place	Pizza Place	Italian Restaurant	Fish & Chips Shop	Fast Food Restaurant	Discount Store	Deli / Bodega
MURRAY DOWNS	Café	Italian Restaurant	Fast Food Restaurant	Discount Store	Sandwich Place	Planetarium	Pizza Place	Fish & Chips Shop	Deli / Bodega
PARINGI	Sandwich Place	Pizza Place	Fish & Chips Shop	Deli / Bodega	Planetarium	Italian Restaurant	Fast Food Restaurant	Discount Store	Café

Dataset after clustering

PostCode	State/Territory	Neighborhood	Latitude	Longitude	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	Cluster Labels
2611	Australian Capital Territory	COOLEMAN, BIMBERI, BRINDABELLA, URIARRA	-35.316175	149.010503	Planetarium	Café	Sandwich Place	Pizza Place	Italian Restaurant	Fish & Chips Shop	Fast Food Restaurant	Discount Store	Deli / Bodega	Beach	0
3500	Victoria	PARINGI	-34.196090	142.142540	Sandwich Place	Pizza Place	Fish & Chips Shop	Deli / Bodega	Planetarium	Italian Restaurant	Fast Food Restaurant	Discount Store	Café	Beach	2
3585	Victoria	MURRAY DOWNS	-35.342298	143.558281	Café	Italian Restaurant	Fast Food Restaurant	Discount Store	Sandwich Place	Planetarium	Pizza Place	Fish & Chips Shop	Deli / Bodega	Beach	2
3644	Victoria	BAROOGA, LALALTY	-35.911559	145.671953	Café	Beach	Sandwich Place	Planetarium	Pizza Place	Italian Restaurant	Fish & Chips Shop	Fast Food Restaurant	Discount Store	Deli / Bodega	1

Visualising Clusters



Result Section

In this section, we analysed the 3 clusters we have found and tried to understand them.

Cluster one

```
#Cluster one  
aus_merged[aus_merged['Cluster Labels'] == 0]
```

PostCode	State/Territory	Neighborhood	Latitude	Longitude	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	Cluster Labels
2611	Australian Capital Territory	COOLEMAN, BIMBERI, BRINDABELLA, URIARRA	-35.316175	149.010503	Planetarium	Café	Sandwich Place	Pizza Place	Italian Restaurant	Fish & Chips Shop	Fast Food Restaurant	Discount Store	Deli / Bodega	Beach	0

Seems to me that the planetarium is the most happening place in this neighborhood followed by food places

Cluster two

```
#Cluster two
aus_merged[aus_merged['Cluster Labels'] == 1]
```

PostCode	State/Territory	Neighborhood	Latitude	Longitude	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	Cluster Labels
3644	Victoria	BAROOGA, LALALTY	-35.911559	145.671953	Café	Beach	Sandwich Place	Planetarium	Pizza Place	Italian Restaurant	Fish & Chips Shop	Fast Food Restaurant	Discount Store	Deli / Bodega	1

This neighborhood favours the beach and tends to have a mix of activity areas and food places in the top 5

Cluster three

```
#Cluster three  
aus_merged[aus_merged['Cluster Labels'] == 2]
```

PostCode	State/Territory	Neighborhood	Latitude	Longitude	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	Cluster Labels
3500	Victoria	PARINGI	-34.196090	142.142540	Sandwich Place	Pizza Place	Fish & Chips Shop	Deli / Bodega	Planetarium	Italian Restaurant	Fast Food Restaurant	Discount Store	Café	Beach	2
3585	Victoria	MURRAY DOWNS	-35.342298	143.558281	Café	Italian Restaurant	Fast Food Restaurant	Discount Store	Sandwich Place	Planetarium	Pizza Place	Fish & Chips Shop	Deli / Bodega	Beach	2

We can clearly see that these neighborhoods favour food as food places are in the top 3



Discussion Section

Since FourSquare API could not return venues for all the Neighborhoods, we were left with limited data to work with. The project can be improved by getting more location data and hopefully more venues returned by FourSquare.

Still with what we've been able to work with, we have managed to find out the food oriented places and the activity oriented places. With more data, we'll be able to get more venue categories like banks, markets, fairs etc



Conclusion Section

After examination of the clusters, we can conclude the following;

1. Neighborhoods: Paringi and Murray Downs are appropriate for foodies as all the top venues are food places
2. The neighborhoods in Australian Capital Territory are useful for anyone who loves the planets and likes to roam around the food places as well
3. The Barooga, Lalalty Neighborhood are for beach lovers

Decisions can me made based on these conclusions



Thankyou