The Battle of Neighbourhoods_Sydney Vs Melbourne

Introduction

For a person like me, to relocate to a new place the following questions pop up in the mind:

1. What would be my rental expenditure?

2. How safe is the area?

3. How well it is connected by public transportation (bus/ train)?

4. How close it is from all the venues?

In quest of the answers, I have done lots of research online and finally came up with a thought to start working on finding solutions by doing some data analysis. As the problem is a very complex one, I am planning on doing it in stages. This article gives a detailed description of my first stage of work.

I chose to tackle two part of the above requirements i.e., connectivity by train and the details of top venues nearby the suburb of choice.

Problem Statement

Part-I

Clustering suburbs of Sydney and Melbourne on the bases of their connectivity to public transport (by train) and comparing how well Sydney and Melbourne are connected by train.

Part-II

Analysing the most common/popular venue details for all suburbs and their comparison.

Data Collection

I am planning to collect the data from the following sources:

- Data related to suburb and train stations from by scraping Wikipedia.
- Venue details from Foursquare develop API.

I wrote an article on medium about how I handling the data collection. It is available on the following link. https://medium.com/@teja37.kota/web-scraping-using-beautifulsoup-made-easy-c2deeac1d231

Scope of the Project

The analysis aiming to be done as a part of the project can be useful for people who are planning to relocate to Melbourne/Sydney.

They can get an information of how far off the train stations are to the suburbs and based on that initial information one can plan to make further decisions.

The analysis also gives an overview of the top venues in the suburbs

Also, from the third part of the project, one can get some insights into the Australian culture based on the details of the most popular venues.

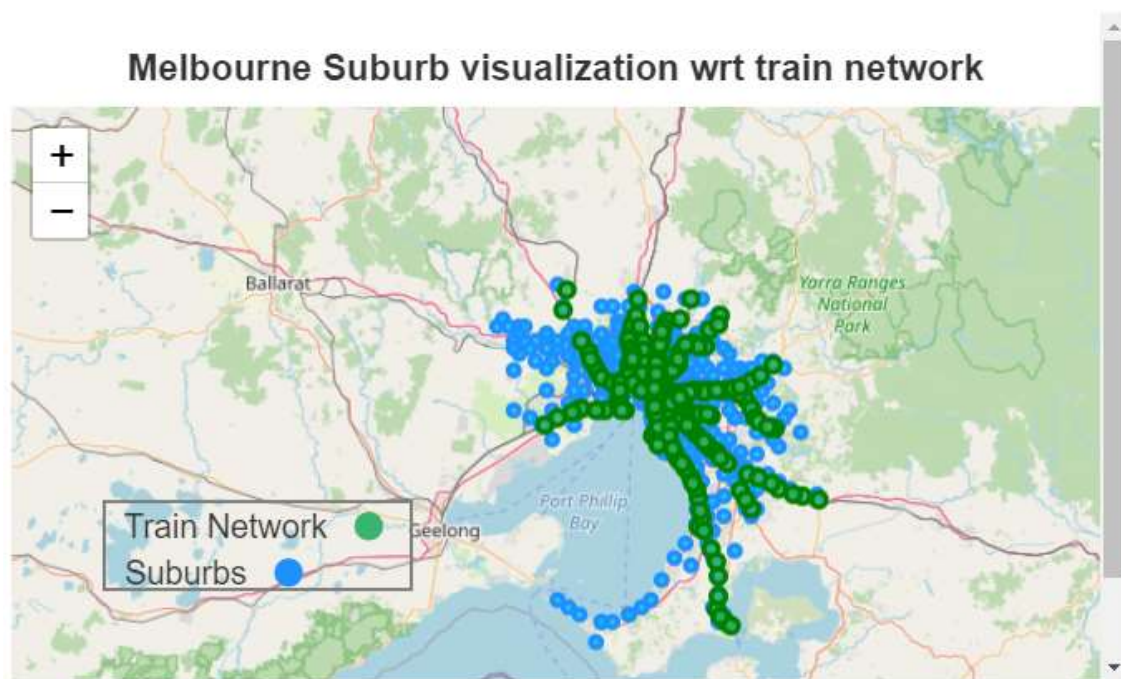# Part-I How well are Melbourne and Sydney connected by Train?

Methodology

Once we are done with the data collection, the most exciting part starts, playing around with data to get some interesting visualizations.
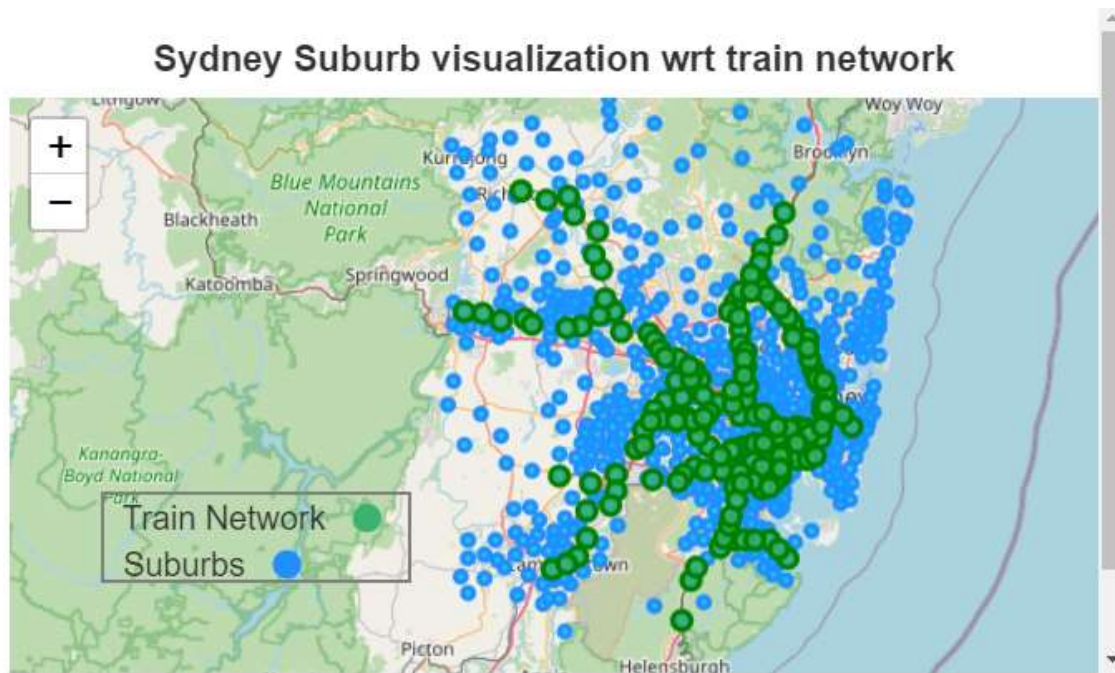
I planned to approach in the following stages:

- **Dropping the stations/suburbs not having the information about their coordinates.**
- **Visualizing the location of suburbs and train stations on the map.**

For this Folium module of Python is used. Here is a map of suburbs with an overlay of the train station network.

Melbourne Train Network Visualization

Sydney Melbourne Train Network Visualization

- **Finding the nearest train station for each suburb.**

The nearest train station is determined by the Euclidean distance between the suburb and stations. We obtain an output with the details of the suburb, its nearest train station and distance between them.

	Suburb	Nearest Train Station	Distance(km)	Latitude	Longitude
0	Abbotsford, Victoria	Victoria Park	0.503514	-37.800278	145.000000
1	Aberfeldie, Victoria	Essendon	1.521641	-37.751111	144.900000
2	Aintree, Victoria	Watergardens	9.595217	-37.716667	144.666667
3	Airport West, Victoria	Glenroy	4.566174	-37.717222	144.867778
4	Albanvale, Victoria	Keilor Plains	3.778670	-37.734444	144.751389
...	...	...	...	...	...
376	Windsor, Victoria	Prahran	0.568560	-37.850278	144.983611
377	Wollert, Victoria	Craigieburn	6.915282	-37.584444	145.018056
378	Wyndham Vale	Werribee	4.177509	-37.883889	144.617778
379	Yallambie, Victoria	Watsonia	1.607158	-37.717500	145.100278
380	Yarraville, Victoria	Yarraville	0.513339	-37.816667	144.883889

The nearest train station information

- **Cluster the suburbs based in the distance from its nearest train station**

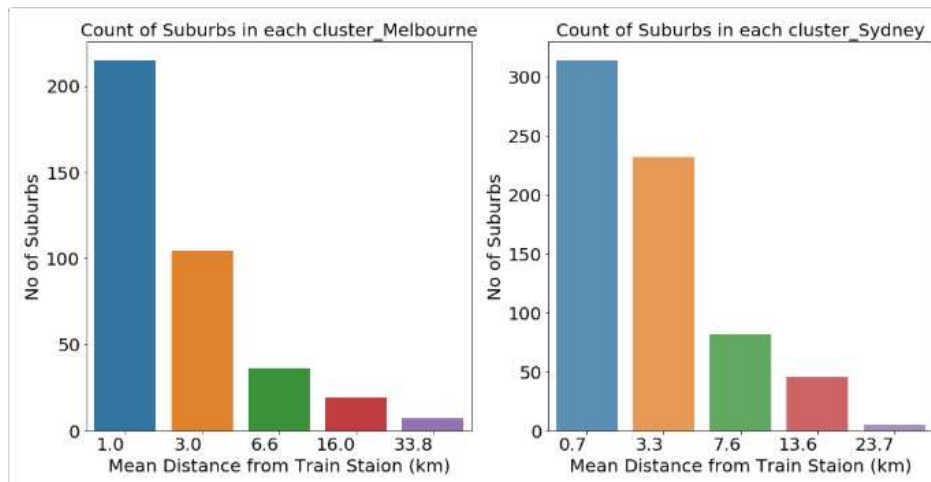I chose to divide the suburbs into five clusters. For this, I used the k-means algorithm. So, now we have a cluster label assigned to each suburb.

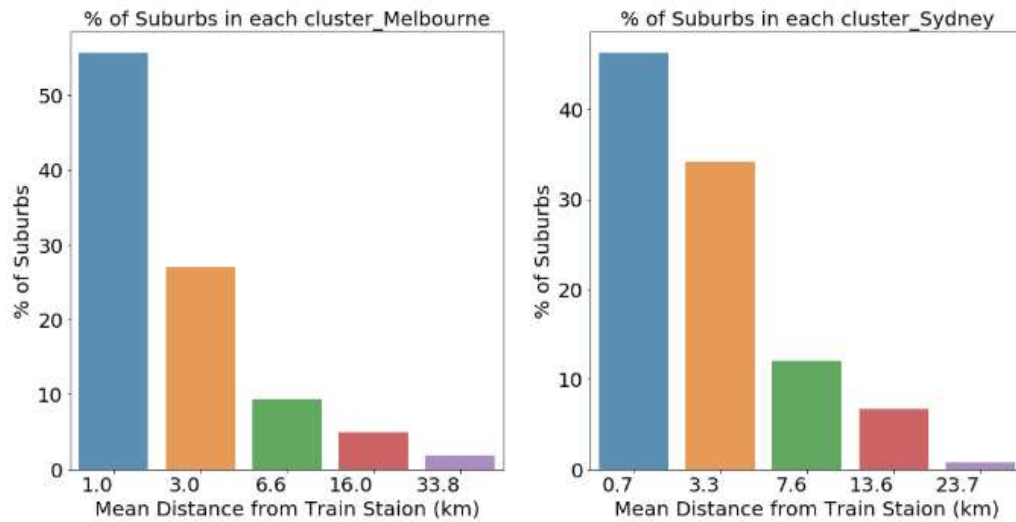|	Suburb	Nearest Train Station	Distance(km)	Cluster_Labels_Dist	Latitude	Longitude
0	Abbotsford, Victoria	Victoria Park	0.503514	0	-37.800278	145.000000
1	Aberfeldie, Victoria	Essendon	1.521641	0	-37.751111	144.900000
2	Aintree, Victoria	Watergardens	9.595217	3	-37.716667	144.666667
3	Airport West, Victoria	Glenroy	4.566174	4	-37.717222	144.867778
4	Albanvale, Victoria	Keilor Plains	3.778670	4	-37.734444	144.751389
...	...	...	...	...	...	...
376	Windsor, Victoria	Prahran	0.568560	0	-37.850278	144.983611
377	Wollert, Victoria	Craigieburn	6.915282	4	-37.584444	145.018056
378	Wyndham Vale	Werribee	4.177509	4	-37.883889	144.617778
379	Yallambie, Victoria	Watsonia	1.607158	0	-37.717500	145.100278
380	Yarraville, Victoria	Yarraville	0.513339	0	-37.816667	144.883889

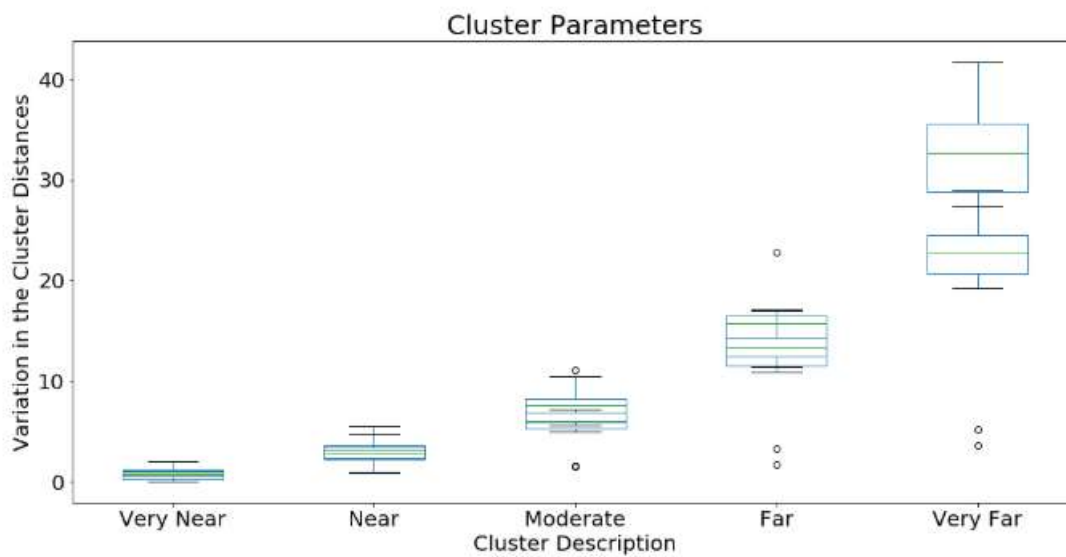Cluster label for the suburbs

Results

Both Sydney and Melbourne are well connected by the train network. Almost 80% of the suburbs have a train station within 5km of distance.

Comparison of the count of suburbs in each cluster (Melbourne& Sydney)

Comparison of the % of suburbs in each cluster (Melbourne& Sydney)

Cluster Parameters

Part-II Analysing the most common/popular venue details for all suburbs and their comparison.

Data Collection

The details of the suburbs of Sydney and Melbourne are scraped from Wikipedia. A previous blog of mine explains the scraping process in detail. Please refer to Web Scraping Using Beautifulsoup made easy.

Along with the suburb information, we need the details of the venues nearby as well. They can be extracted from Foursquare Developer API.

Details of 3122 venues, 313 venues categories in Melbourne and 5241 venues, 380 venues categories in Sydney were extracted and analysed.

venues in Melbourne:

	Suburb	Suburb Latitude	Suburb Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Abbotsford, Victoria	-37.800278	145.000000	Lentil As Anything	-37.802724	145.003507	Vegetarian / Vegan Restaurant
1	Abbotsford, Victoria	-37.800278	145.000000	Abbotsford Convent	-37.802574	145.004410	Cultural Center
2	Abbotsford, Victoria	-37.800278	145.000000	Retreat Hotel	-37.801126	144.997548	Pub
3	Abbotsford, Victoria	-37.800278	145.000000	Slow Food Market	-37.802481	145.003597	Farmers Market
4	Abbotsford, Victoria	-37.800278	145.000000	The Farm Cafe	-37.802940	145.005706	Café
...	...	...	...	...	...	...	...
4074	Wyndham Vale	-37.883889	144.617778	Aldi	-37.887292	144.607565	Department Store
4075	Wyndham Vale	-37.883889	144.617778	Presidents Park Softball Grounds	-37.882215	144.633154	Baseball Field
4076	Wyndham Vale	-37.883889	144.617778	Beautiful Cooking	-37.891832	144.630117	Asian Restaurant
4077	Wyndham Vale	-37.883889	144.617778	Wyndham Vale Station	-37.872706	144.608604	Train Station
4078	Wyndham Vale	-37.883889	144.617778	Manor Lakes Football Club	-37.874627	144.604154	Athletics & Sports

Details of Venues in Melbourne

Venues in Sydney:

	Suburb	Suburb Latitude	Suburb Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Abbotsbury	-33.869285	150.866703	Calmsley Hill City Farm	-33.868848	150.858808	Farm
1	Abbotsbury	-33.869285	150.866703	Abbotsbury Shops	-33.869554	150.865294	Convenience Store
2	Abbotsbury	-33.869285	150.866703	Stockdale Reserve	-33.871098	150.868060	Park
3	Abbotsbury	-33.869285	150.866703	817 bus stop	-33.869773	150.870122	Bus Station
4	Abbotsford	-33.850553	151.129759	The Cove Dining Co	-33.850933	151.129015	Café
...	...	...	...	...	...	...	...
6501	South Granville	-33.857778	151.010556	Aldadoun	-33.857936	151.010928	Middle Eastern Restaurant
6502	South Granville	-33.857778	151.010556	El Dadoun Butchery	-33.858272	151.010685	Middle Eastern Restaurant
6503	South Granville	-33.857778	151.010556	Ray Marshall Reserve	-33.861963	151.012686	Soccer Field
6504	South Granville	-33.857778	151.010556	Fauna Reserve	-33.858799	151.017160	Garden
6505	South Granville	-33.857778	151.010556	Aviary	-33.859701	151.017473	Park

Details of Venues in Sydney

Methodology

Many insights can be drawn from the collected datasets. However, for this project, my aim is to figure out the top ten venues in all suburbs and also compare the top ten venues categories of Sydney and Melbourne.

To obtain the data we need, One hot encoder is applied to the data sets, resulting in a dataframe with columns equivalent to the no of unique venue categories.

The count of each venue category is estimated for all the suburbs and the top 10 venue categories are then extracted.

Top 10 venues in Melbourne Suburbs

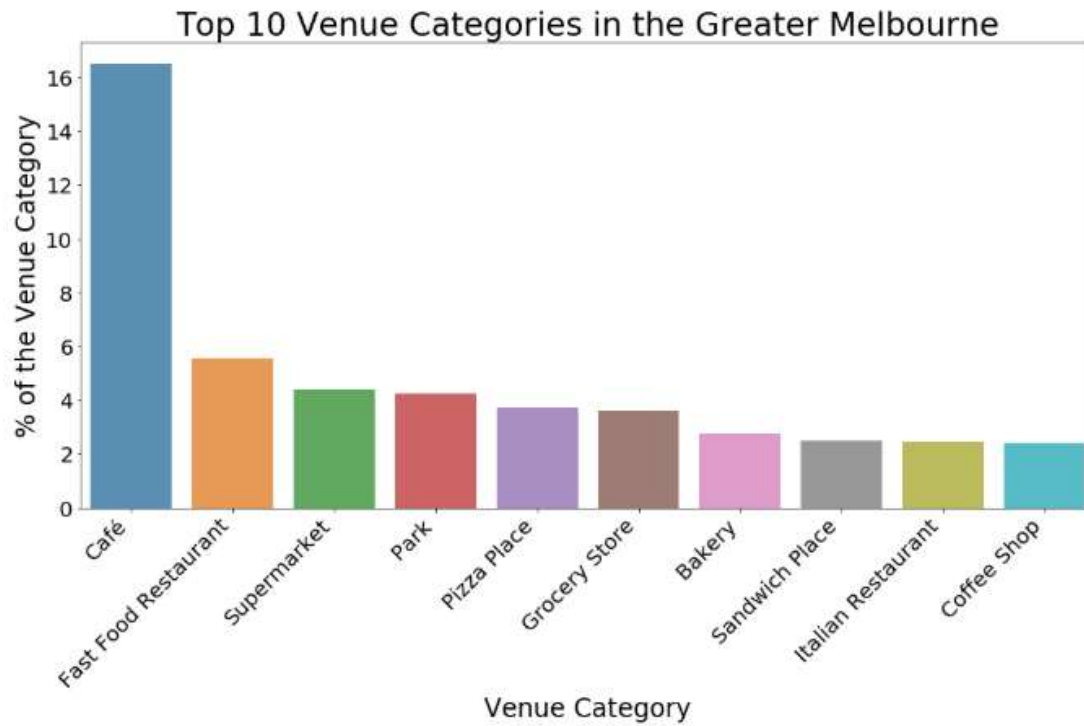|	Suburb	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Abbotsford, Victoria	Café	Bus Stop	Pub	Farmers Market	Farm	Gastropub	Scenic Lookout	Greek Restaurant	Grocery Store	Thrift / Vintage Store
1	Aberfeldie, Victoria	Café	Grocery Store	Spa	Food & Drink Shop	Coffee Shop	Zoo Exhibit	Field	Film Studio	Fish & Chips Shop	Fish Market
2	Aintree, Victoria	Furniture / Home Store	Playground	Fried Chicken Joint	French Restaurant	Fast Food Restaurant	Field	Film Studio	Fish & Chips Shop	Fish Market	Fishing Spot
3	Airport West, Victoria	Dessert Shop	Rental Car Location	Café	Farmers Market	Forest	Football Stadium	Food Truck	Food Court	Food & Drink Shop	Food
4	Albanvale, Victoria	Spa	Sushi Restaurant	Thai Restaurant	Hotel	Zoo Exhibit	Flea Market	Fast Food Restaurant	Field	Film Studio	Fish & Chips Shop
...	...	...	...	...	...	...	...	...	...	...	...

Top 10 Venue details of Suburb in Melbourne

Top 10 venues in Sydney Suburbs

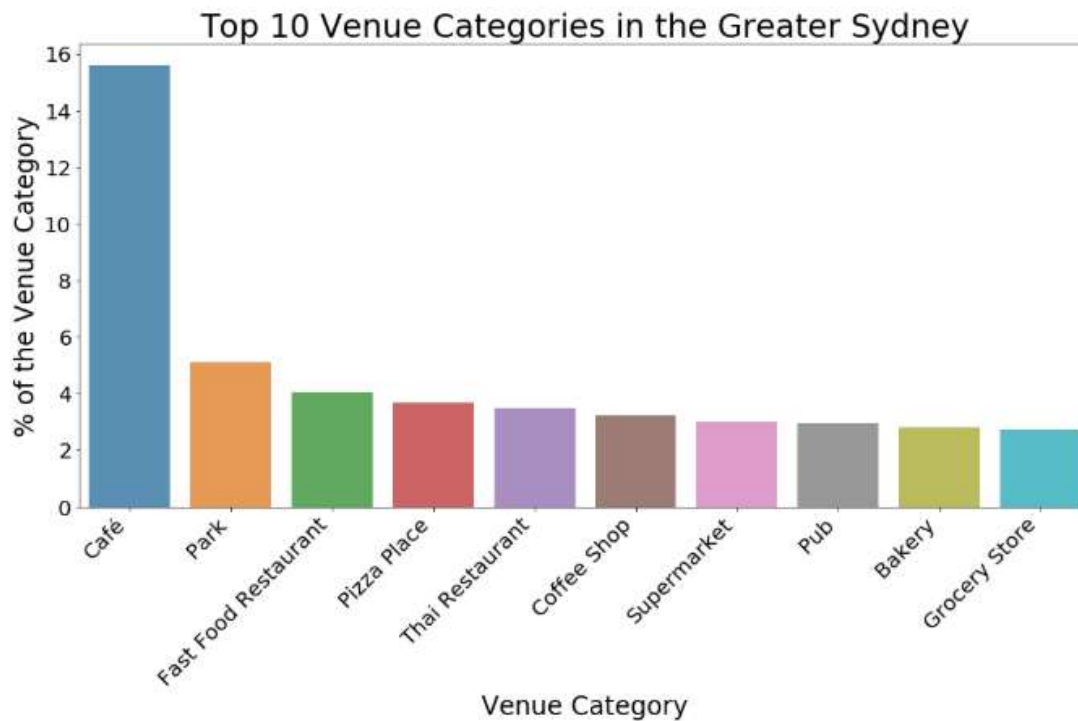|	Suburb	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Jordan Springs	Gym	Lake	Café	Department Store	Auto Garage	Fishing Store	Food	Dutch Restaurant	Eastern European Restaurant	Egyptian Restaurant
1	South Granville	Middle Eastern Restaurant	Post Office	Soccer Field	Park	Garden	Fishing Spot	Fish Market	Eastern European Restaurant	Flower Shop	Egyptian Restaurant
2	Abbotsbury	Park	Convenience Store	Bus Station	Farm	Farmers Market	Eastern European Restaurant	Egyptian Restaurant	Electronics Store	Event Service	Event Space
3	Abbotsford	Café	Pizza Place	Thai Restaurant	Park	Construction & Landscaping	Wine Shop	Costume Shop	Dumpling Restaurant	Eastern European Restaurant	Egyptian Restaurant
4	Agnes Banks	Park	Nature Preserve	Other Repair Shop	Farm	Farmers Market	Eastern European Restaurant	Egyptian Restaurant	Electronics Store	Event Service	Event Space
...	...	...	...	...	...	...	...	...	...	...	...

Top 10 Venue details of Suburb in Sydney

Results

Cafes are the most popular venue categories in Sydney and Melbourne. Among the data analysed more than 15% are cafes.

Top 10 venue categories in Melbourne

Top 10 venue categories in Sydney

	Melbourne	Count Melbourne	% Melbourne	Sydney	Count Sydney	% Sydney
0	Café	515	16.495836	Café	817	15.588628
1	Fast Food Restaurant	174	5.573350	Park	268	5.113528
2	Supermarket	137	4.388213	Fast Food Restaurant	212	4.045030
3	Park	133	4.260090	Pizza Place	192	3.663423
4	Pizza Place	117	3.747598	Thai Restaurant	182	3.472620
5	Grocery Store	113	3.619475	Coffee Shop	169	3.224575
6	Bakery	86	2.754644	Supermarket	158	3.014692
7	Sandwich Place	78	2.498398	Pub	156	2.976531
8	Italian Restaurant	77	2.466368	Bakery	148	2.823889
9	Coffee Shop	75	2.402306	Grocery Store	144	2.747567

Count and % of the top 10 venue categories in Melbourne and Sydney