

Exploring The Emerald City(Sydney)

Introduction:

1.1. BackGround :

Sydney is the capital of New South Wales In Australia.It is also Know as Emerald City and Harbour City Of Australia.It is well know for it's tourist places.It has many tourist attractions like Beaches,Harbours,Bays.

Sydney Opera House is one of the most visited tourist spot in sydney because of it's unique structure.

1.2 Problem:

Sydney is one of the largest city.For visitors it is difficult to find the venues that are nearby in the city.For that in this project iam using the foursquare api data to get the venues that are available in the city,and clustering them based on the category. So that the visitors can easily find restaurants, bars, tourist spots,coffe shops.

1.3 Interest:

People who are willing to visit sydney this project will provides details of all the venues on interactive maps.

Data Acquistion and Cleaning:

2.1 Data Sources:

In this project,I extract the data from foursquare api which provides the venue details of a particular city.The data that is received from foursquare api is in json format.

2.2 Data Cleaning:

we need to extract the data from json file into a data frame by json_normalise method.In that json files it contains many details like location,city,country,name,category etc.

We need only name of the restaurant,category,latitude and longitude coordinates.After getting the data into a data frame we need to check for any null values if any drop those rows.

2.3 Feature Selection:

In this our main aim is to cluster the venues based on the category of the venues.since the category has more than 52 unique values we change the category of all the data by looping the category column and create a new column category type.

Before adding category_type column:

Out[80]:

	name	lat	lng	category
0	Sydney Opera House - Concert Hall	-33.856595	151.215058	Concert Hall
1	Drama Theatre	-33.856863	151.214759	Theater
2	Sydney Opera House	-33.857260	151.215040	Opera House
3	Sydney Opera House - Studio	-33.856991	151.214482	Theater
4	Sydney Opera House - Playhouse	-33.857518	151.214559	Theater
5	Opera Kitchen	-33.857972	151.213989	Café
6	Opera Bar	-33.858409	151.213976	Cocktail Bar
7	Joan Sutherland Theatre	-33.857514	151.214286	Concert Hall
8	Sydney Opera House Forecourt	-33.858184	151.214619	Plaza
9	Sydney Harbour Bridge	-33.852289	151.210681	Bridge
10	Bennelong Restaurant	-33.857564	151.214794	Australian Restaurant
11	Kirribilli Lookout	-33.849323	151.213298	Scenic Lookout
12	Harbour Bridge Pylon Lookout	-33.854580	151.209492	Scenic Lookout
13	Mrs. Macquarie's Point	-33.859395	151.222204	Scenic Lookout
14	The Opera House to the Botanic Gardens Walk	-33.860914	151.213221	Trail
15	Park Hyatt Sydney	-33.856023	151.209225	Hotel
16	Mrs Macquarie's Chair	-33.859534	151.222539	Monument / Landmark
17	Aria Restaurant	-33.859198	151.213261	Australian Restaurant
18	Quay	-33.857445	151.210155	Australian Restaurant
19	Portside Sydney	-33.857050	151.214650	Australian Restaurant

After adding category_type column:

Out[87]:

	name	lat	lng	category	category type
0	Sydney Opera House - Concert Hall	-33.856595	151.215058	Concert Hall	Theatre
1	Drama Theatre	-33.856863	151.214759	Theater	Theatre
2	Sydney Opera House	-33.857260	151.215040	Opera House	Tourist Spots
3	Sydney Opera House - Studio	-33.856991	151.214482	Theater	Theatre
4	Sydney Opera House - Playhouse	-33.857518	151.214559	Theater	Theatre
5	Opera Kitchen	-33.857972	151.213989	Café	coffee shop
6	Opera Bar	-33.858409	151.213976	Cocktail Bar	Bar
7	Joan Sutherland Theatre	-33.857514	151.214286	Concert Hall	Theatre
8	Sydney Opera House Forecourt	-33.858184	151.214619	Plaza	Retail Store and Markets
9	Sydney Harbour Bridge	-33.852289	151.210681	Bridge	Tourist Spots
10	Bennelong Restaurant	-33.857564	151.214794	Australian Restaurant	Restaurant
11	Kirribilli Lookout	-33.849323	151.213298	Scenic Lookout	Tourist Spots
12	Harbour Bridge Pylon Lookout	-33.854580	151.209492	Scenic Lookout	Tourist Spots
13	Mrs. Macquarie's Point	-33.859395	151.222204	Scenic Lookout	Tourist Spots
14	The Opera House to the Botanic Gardens Walk	-33.860914	151.213221	Trail	Tourist Spots
15	Park Hyatt Sydney	-33.856023	151.209225	Hotel	Restaurant
16	Mrs Macquarie's Chair	-33.859534	151.222539	Monument / Landmark	Tourist Spots
17	Aria Restaurant	-33.859198	151.213261	Australian Restaurant	Restaurant
18	Quay	-33.857445	151.210155	Australian Restaurant	Restaurant
19	Portside Sydney	-33.857050	151.214650	Australian Restaurant	Restaurant



Now you can see the difference after adding category_type column.

Clustering the Venues into different clusters:

The next step is cluster the venues into different types using kmeans clustering algorithm.

We have total 11 unique categories and so I cluster them into 11 clusters.

Cluster1:

It contains all the coffee shops near by the city.

```
[30]: cluster1
```

```
ut[30]:
```

	name	lat	lng	Cluster_Labels	category	category type
5	Opera Kitchen	-33.857972	151.213989	0	Café	coffee shop
21	The Tea Cosy	-33.857413	151.208561	0	Café	coffee shop
44	The Fine Food Store	-33.858546	151.208633	0	Café	coffee shop
48	Café Opera	-33.863089	151.212463	0	Café	coffee shop
50	Cabrito Coffee Traders	-33.862516	151.209324	0	Café	coffee shop
57	Celsius Coffee Co.	-33.849403	151.220256	0	Café	coffee shop
67	Bertoni	-33.865167	151.211522	0	Café	coffee shop
73	Creperie Suzette	-33.859534	151.207994	0	Café	coffee shop
87	Marlowe's Way	-33.863877	151.208744	0	Café	coffee shop
89	Georgie Boy's Coffee Co.	-33.862403	151.207591	0	Coffee Shop	coffee shop

Cluster2:It contains all the Tourist Spots.

Out[32]:

	name	lat	lng	Cluster_Labels	category	category type
2	Sydney Opera House	-33.857260	151.215040	1	Opera House	Tourist Spots
9	Sydney Harbour Bridge	-33.852289	151.210681	1	Bridge	Tourist Spots
11	Kirribilli Lookout	-33.849323	151.213298	1	Scenic Lookout	Tourist Spots
12	Harbour Bridge Pylon Lookout	-33.854580	151.209492	1	Scenic Lookout	Tourist Spots
13	Mrs. Macquarie's Point	-33.859395	151.222204	1	Scenic Lookout	Tourist Spots
14	The Opera House to the Botanic Gardens Walk	-33.860914	151.213221	1	Trail	Tourist Spots
16	Mrs Macquarie's Chair	-33.859534	151.222539	1	Monument / Landmark	Tourist Spots
20	Dawes Point	-33.855194	151.209820	1	Scenic Lookout	Tourist Spots
22	Government House	-33.860616	151.214903	1	History Museum	Tourist Spots
25	BridgeClimb Sydney	-33.857518	151.207832	1	Tour Provider	Tourist Spots
29	Museum of Contemporary Art	-33.859859	151.209018	1	Art Museum	Tourist Spots
32	Royal Botanic Garden	-33.864144	151.215992	1	Botanical Garden	Tourist Spots
33	Main Pond	-33.863542	151.217032	1	Garden	Tourist Spots
37	Cahill Walk	-33.852365	151.210531	1	Trail	Tourist Spots
38	Cahill Expressway Lookout	-33.861265	151.211277	1	Scenic Lookout	Tourist Spots
49	The Big Tree	-33.862583	151.217045	1	Tree	Tourist Spots
59	Lady Gowrie Lookout	-33.850945	151.219628	1	Scenic Lookout	Tourist Spots
72	Pier 2/3	-33.855009	151.207213	1	Harbor / Marina	Tourist Spots
76	Sydney Observatory	-33.859534	151.204643	1	Planetarium	Tourist Spots
78	Handa Opera On Sydney Harbour	-33.861466	151.220863	1	Opera House	Tourist Spots
91	Kurraba Point Reserve	-33.845040	151.223208	1	Scenic Lookout	Tourist Spots
94	Wendy Whiteley's Secret Garden	-33.843822	151.208130	1	Garden	Tourist Spots

Cluster 3:It contains all the parks,pool information.

Out[34]:

	name	lat	lng	Cluster_Labels	category	category type
27	Bradfield Park	-33.848728	151.212132	2	Park	Entertainment
31	Dr Mary Booth Park & Lookout	-33.850418	151.215772	2	Park	Entertainment
43	Rooftop Pool (Park Hyatt)	-33.855876	151.209386	2	Pool	Entertainment
45	Andrew (Boy) Charlton Pool	-33.863664	151.221581	2	Pool	Entertainment
52	Observatory Hill	-33.859125	151.204977	2	Park	Entertainment
60	North Sydney Olympic Pool	-33.849320	151.210886	2	Pool	Entertainment
70	Blues Point Reserve	-33.848932	151.204026	2	Park	Entertainment
83	Barangaroo Reserve	-33.857052	151.201100	2	Park	Entertainment
93	The Domain	-33.867868	151.215370	2	Park	Entertainment
99	Luna Park	-33.848722	151.210515	2	Theme Park	Entertainment

Cluster 4:It contains all the restaurants information.

```
In [36]: cluster4
```

```
Out[36]:
```

	name	lat	lng	Cluster_Labels	category	category type
10	Bennelong Restaurant	-33.857564	151.214794	3	Australian Restaurant	Restaurant
15	Park Hyatt Sydney	-33.856023	151.209225	3	Hotel	Restaurant
17	Aria Restaurant	-33.859198	151.213261	3	Australian Restaurant	Restaurant
18	Quay	-33.857445	151.210155	3	Australian Restaurant	Restaurant
19	Portside Sydney	-33.857050	151.214650	3	Australian Restaurant	Restaurant
23	Yukis At The Quay	-33.858409	151.210287	3	Japanese Restaurant	Restaurant
24	6 Head 1788	-33.856670	151.209150	3	Australian Restaurant	Restaurant
47	InterContinental Sydney	-33.863076	151.212066	3	Hotel	Restaurant
54	Chat Thai ชัดไทย (Chat Thai)	-33.861872	151.209958	3	Thai Restaurant	Restaurant
55	Café Sydney	-33.862217	151.210934	3	Australian Restaurant	Restaurant
56	La Capannina Restaurant	-33.847297	151.212090	3	Italian Restaurant	Restaurant
66	Shangri-La Hotel	-33.861141	151.206460	3	Hotel	Restaurant
68	The East Chinese Restaurant	-33.859303	151.212921	3	Chinese Restaurant	Restaurant
69	The Spice Room	-33.861757	151.212293	3	Indian Restaurant	Restaurant
82	Restaurant Hubert	-33.865289	151.210303	3	French Restaurant	Restaurant
86	Lotus Dumpling Bar	-33.856440	151.206390	3	Asian Restaurant	Restaurant
88	Fish at the Rocks	-33.858673	151.203500	3	Seafood Restaurant	Restaurant
90	The Langham Hotel Sydney	-33.860517	151.203437	3	Hotel	Restaurant
95	Mr. Wong	-33.864113	151.207970	3	Chinese Restaurant	Restaurant
96	Ainoya (藍の屋)	-33.846227	151.213450	3	Japanese Restaurant	Restaurant

Cluster 5:It contain bars nearby the city.


```
In [37]: cluster5=sydney.loc[sydney['Cluster_Labels']==4]
cluster5
```

Out[37]:

	name	lat	lng	Cluster_Labels	category	category type
6	Opera Bar	-33.858409	151.213976	4	Cocktail Bar	Bar
26	The Bar at Park Hyatt Sydney	-33.856937	151.209319	4	Hotel Bar	Bar
30	The Glenmore Rooftop Hotel	-33.858632	151.207464	4	Pub	Bar
34	The Squire's Landing	-33.857346	151.210219	4	Brewery	Bar
36	Bulletin Place	-33.862344	151.209339	4	Cocktail Bar	Bar
39	Garden Bar by the Corner House	-33.857113	151.214437	4	Cocktail Bar	Bar
58	The Cortile	-33.862977	151.212463	4	Hotel Bar	Bar
62	Club Intercontinental	-33.862820	151.212489	4	Hotel Bar	Bar
64	Blu Bar On 36	-33.861067	151.206361	4	Hotel Bar	Bar
71	Busby's Bar	-33.862598	151.220418	4	Beer Garden	Bar
74	Harts Pub	-33.861870	151.206314	4	Pub	Bar
75	Lord Nelson Brewery Hotel	-33.858403	151.203548	4	Brewery	Bar
76	Palmer & Co.	-33.864208	151.208340	4	Speakeasy	Bar
81	Palisade Hotel	-33.857979	151.202264	4	Pub	Bar
97	Hacienda	-33.860862	151.212290	4	Cocktail Bar	Bar

click to scroll output; double click to hide

Cluster 6:It contains all Theater that are available nearby the city.


```
In [38]: cluster6=sydney.loc[sydney['cluster_Labels']==5]
cluster6
```

Out[38]:

	name	lat	lng	Cluster_Labels	category	category type
0	Sydney Opera House - Concert Hall	-33.856595	151.215058	5	Concert Hall	Theatre
1	Drama Theatre	-33.856863	151.214759	5	Theater	Theatre
3	Sydney Opera House - Studio	-33.856991	151.214482	5	Theater	Theatre
4	Sydney Opera House - Playhouse	-33.857518	151.214559	5	Theater	Theatre
7	Joan Sutherland Theatre	-33.857514	151.214286	5	Concert Hall	Theatre
41	St. George OpenAir Cinema	-33.861503	151.220864	5	Movie Theater	Theatre
46	Dendy Cinemas	-33.859505	151.212910	5	Indie Movie Theater	Theatre
63	Roslyn Packer Theatre	-33.857019	151.204947	5	Theater	Theatre

Cluster 7:It contains all the fast food centers.

```
In [39]: cluster7=sydney.loc[sydney['cluster_Labels']==6]
cluster7
```

Out[39]:

	name	lat	lng	Cluster_Labels	category	category type
28	Ribs & Burgers	-33.857753	151.209045	6	Burger Joint	Fast Food
61	Gateway Foodcourt	-33.862051	151.209879	6	Food Court	Fast Food
65	Bar Luca	-33.864078	151.211626	6	Burger Joint	Fast Food
98	Frankie's Pizza	-33.866013	151.209588	6	Pizza Place	Fast Food

Cluster 8:It contains yoga studio and library in the city.

```
In [40]: cluster8=sydney.loc[sydney['Cluster_Labels']==7]
cluster8
```

Out[40]:

	name	lat	lng	Cluster_Labels	category	category type
42	Sydney Theatre Company	-33.857028	151.204938	7	Performing Arts Venue	Fitness
77	Sydney Conservatorium of Music	-33.863457	151.213779	7	Music School	Fitness
84	Yoga Beyond	-33.863768	151.221375	7	Yoga Studio	Fitness
85	State Library of New South Wales	-33.866693	151.212483	7	Library	Fitness

Cluster 9:It contains the dessert shops.

```
In [41]: cluster9=sydney.loc[sydney['Cluster_Labels']==8]
cluster9
```

Out[41]:

	name	lat	lng	Cluster_Labels	category	category type
35	La Renaissance Café Pâtisserie	-33.859260	151.208428	8	Bakery	Desserts
40	Gelato Messina	-33.862090	151.209860	8	Ice Cream Shop	Desserts
80	Gelatissimo	-33.861283	151.211651	8	Ice Cream Shop	Desserts

Cluster 10:It contain all the markets.

```
In [42]: cluster10=sydney.loc[sydney['Cluster_Labels']==9]
cluster10
```

Out[42]:

	name	lat	lng	Cluster_Labels	category	category type
8	Sydney Opera House Forecourt	-33.858184	151.214619	9	Plaza	Retail Store and Markets
92	Rockpool Bar & Grill	-33.866206	151.210120	9	Steakhouse	Retail Store and Markets

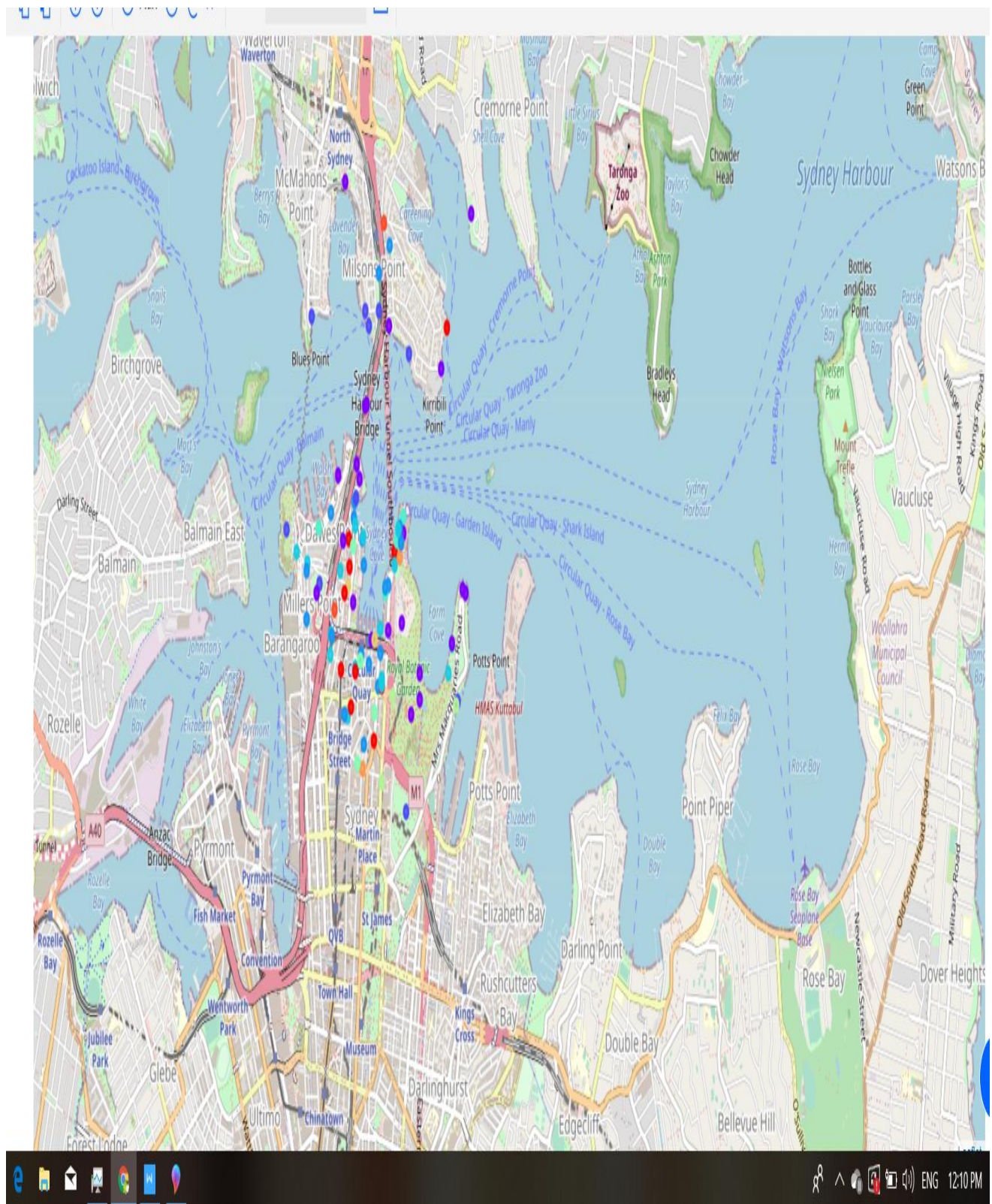
Cluster 11:It contains about accomodation venues nearby the city.

```
In [43]: cluster11=sydney.loc[sydney['Cluster_Labels']==10]
cluster11
```

Out[43]:

	name	lat	lng	Cluster_Labels	category	category type
51	Sydney Harbour YHA	-33.860128	151.206844	10	Hostel	Accomodation
53	The Botanist	-33.845386	151.212772	10	Lounge	Accomodation

Finally plotting all the clusters on folium map.



Conclusion:

In this project i explore the sydney using foure square api within 2km range.Next cluster the venues into different categories and finally plotting them on the folium map.This project is useful to find the venues that are avilable.Each cluster contains the different venues and it is to find the venue that we are looking by simply looking into the clusters.

Future Directions:

In future we can add directions to venues from our loaction and also the distance from our location to all the venues.