



BATTLE OF THE NEIGHBOURHOODS

***Finding the best location to open a new Coffee-house
in Sydney***



Sapna Yadav

CONTENTS

- BUSINESS CASE
- DATA DESCRIPTION
- METHODOLOGY
- RESULTS
- CONCLUSION



BUSINESS CASE

- One of the popular retail coffeehouse chain company is planning to open a new coffee shop in **Sydney, NSW**.
- It is trying to decide which suburb of Sydney would be the best one to locate the Coffee shop.
- The company wants to prefer those location where coffee shops are the **most visited venue** so that will be good opportunity for them to open a coffee shop in that area.
- This project is particularly useful to investors, business owners looking to expand their business in one of the most liveable cities of Australia



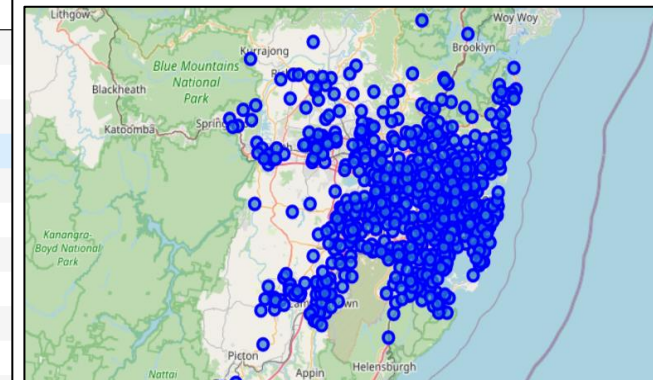
DATA DESCRIPTION

- Data Requirements
 - List of suburbs in Sydney
 - Suburb-wise list of venues
- Data Sources
 - Intosydneydirectory.com
 - Four Square Location API
- Data Processing
 - Data cleansing
 - Structuring data to appropriate format

Sydney			into Sydney
Sydney Postcodes - Sydney Suburbs			Search Engine Add Your Business Business Websites Terms Of Service Contact
Abbotsbury	NSW	2176	
Abbotsford	NSW	2046	
Agnes Banks	NSW	2753	
Airds	NSW	2560	
Alexandria	NSW	2015	
Alfords Point	NSW	2234	
Allambie Heights	NSW	2100	
Allawah	NSW	2218	
Ambarvale	NSW	2560	
Annandale	NSW	2038	
Annangrove	NSW	2156	

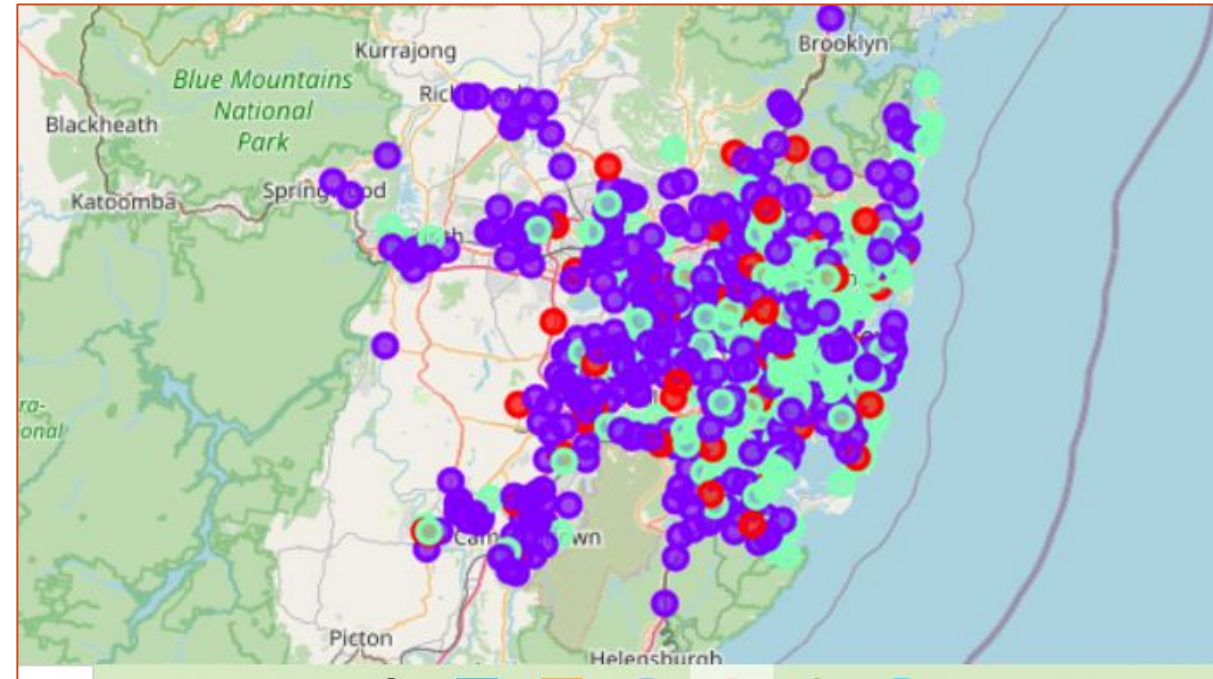


Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue
Abbotsbury	4	4	4
Abbotsford	7	7	7
Airds	4	4	4
Alexandria	31	31	31
Alfords Point	5	5	5
Allambie Heights	5	5	5
Allawah	4	4	4
Ambarvale	7	7	7
Annandale	25	25	25
Appin	3	3	3



METHODOLOGY

- Data Scrapping Technique
 - Web scraping using pandas library to extract the required data
- Foursquare.com
 - Retrieving venue information using API
- Data Exploration
 - Dataset to display the top venue for each neighbourhood
- K-Means Machine Learning
 - Segment and cluster similar neighbourhoods
 - Group them based on most common venues



	Neighborhood	1st Most Common Venue
0	Abbotsbury	Convenience Store
1	Abbotsford	Café
2	Airds	Pub
3	Alexandria	Café
4	Alfords Point	Candy Store



RESULTS

- Clustering Similar Neighbourhoods
 - Grouping based on 1st preference of venue
 - Topmost common venue

	Neighborhood	Postcode	Latitude	Longitude	1st Most Common Venue	Labels
11	Appin	2560	-34.084740	150.806659	Business Service	0.0
14	Arndell Park	2148	-33.788370	150.879830	Outdoors & Recreation	0.0
22	Austral	2179	-33.926835	150.808311	Park	0.0
28	Balgowlah Heights	2093	-33.804960	151.262050	Park	0.0
30	Balmain East	2041	-33.857680	151.191370	Park	0.0
39	Bardia	2565	-33.977234	150.861874	Park	0.0
41	Bardwell Valley	2207	-33.937130	151.133050	Park	0.0
76	Blakehurst	2221	-33.987970	151.111490	Sporting Goods Shop	0.0
105	Cabarita	2137	-33.845680	151.115550	Park	0.0
135	Cawdor	2570	-34.059512	150.689893	Park	0.0
141	Chatswood West	2067	-33.792570	151.159490	Park	0.0
167	Condell Park	2200	-33.920710	151.005100	Park	0.0

Park	43
Café	2
Grocery Store	2
Construction & Landscaping	1
Sporting Goods Shop	1
Auto Garage	1
Outdoors & Recreation	1
Scenic Lookout	1
Gym	1
Business Service	1
Name: 1st Most Common Venue, dtype: int64	

Cluster 1

Fast Food Restaurant	25
Café	19
Convenience Store	16
Pub	15
Platform	12
Coffee Shop	11
Grocery Store	10
Thai Restaurant	10
Supermarket	10
Pizza Place	10
Bakery	10
Shopping Mall	8
Liquor Store	8
Playground	7
Indian Restaurant	6
Golf Course	6
Gym	6
Business Service	6
Chinese Restaurant	6
Lebanese Restaurant	5
Electronics Store	5

Cluster 2

Café	176
Grocery Store	7
Convenience Store	6
Bakery	6
Pub	3
Liquor Store	2
Japanese Restaurant	2
Gym	2
Indian Restaurant	2
Bar	2
Garden Center	1
Scenic Lookout	1
Paintball Field	1
Eastern European Restaurant	1
Print Shop	1
Golf Course	1
Burger Joint	1
Cosmetics Shop	1
Fried Chicken Joint	1
Athletics & Sports	1
Beach	1
Gas Station	1

Cluster 3



CONCLUSION

- **Cluster 1**, which comprises of 7.8% of total suburbs in Sydney, the most common visited venue in this cluster is **Park** which forms 72% of the first preference in this cluster. The percentage to visit café as first preference is 3.7%
- **Cluster 2** which comprises of 49.7% of total suburbs in Sydney, the most common visited venue in this cluster is **Fast Food restaurant** which is 7.28% of this cluster. The percentage to visit café as first preference is 5.53%
- **Cluster 3** which comprises of 32.9% of total suburbs in Sydney, the percentage to visit the Café as first preference is 77.53% for this cluster
- We can conclude that this analysis might allow Cluster 3 suburb locations to company to open a new coffeehouse
- In future we can explore the population of each suburb. Also, to explore the distance from the main CBD area of Sydney to that suburb would most likely have impact here also.

