

Introduction

Bakeries are a popular type of foodservice establishment, allowing people to express creativity while also serving a unique market. Starting a bakery business presents many unique challenges. Just like the other business, bakery is a serious investment. This project presents deeper knowledge prior to opening a bakery, and investigate similarities and differences between suburbs in Sydney.

Business Problem

This project answers the following question:

Where is the best location to open a bakery in Sydney?

Data

1. Suburbs in Sydney, Australia:

<https://www.kaggle.com/ssk27997/suburbs-in-sydney-australia>, containing the list of suburbs in Sydney along with their latitude and longitude.

2. Foursquare Places API

Used to search all venues in a suburb.

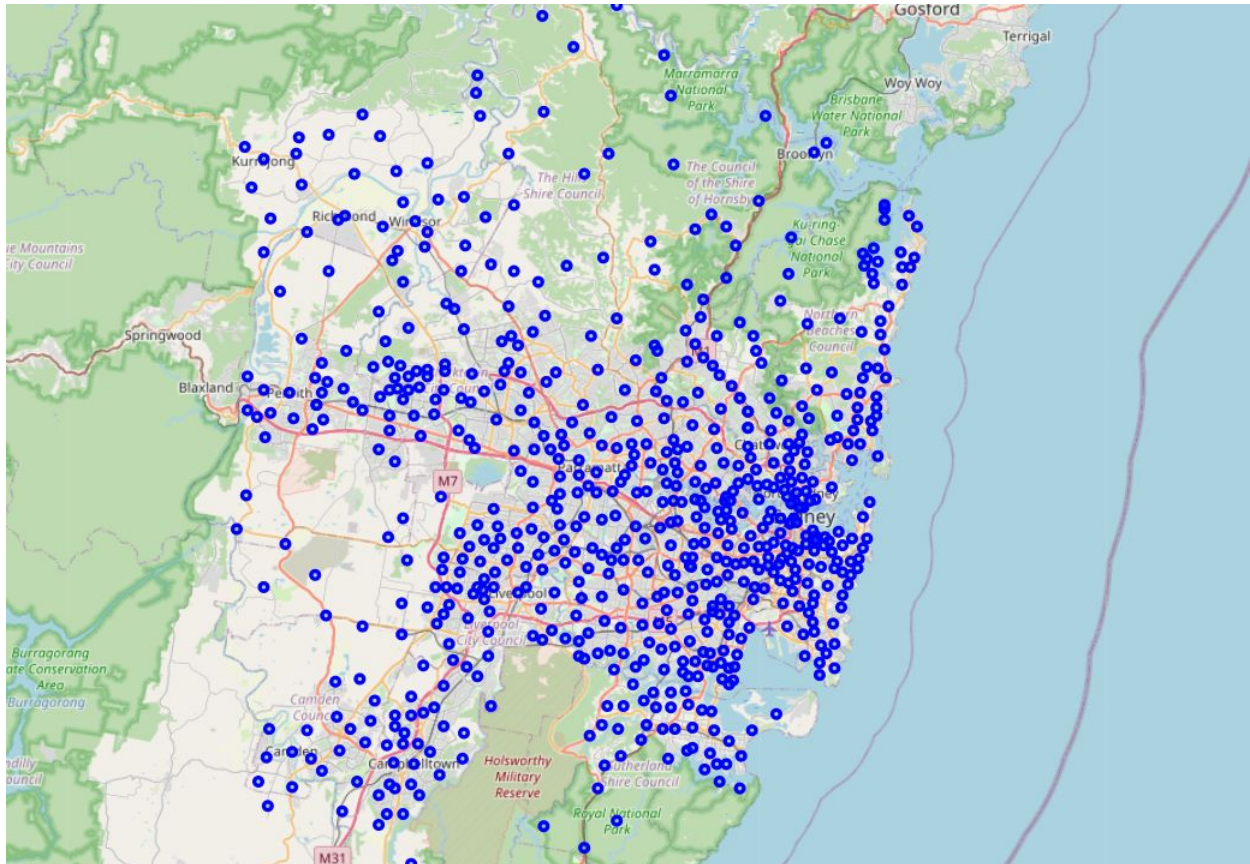
Methodology

1. Explore the data
2. Use Foursquare Places API to obtain venue information from each suburb
3. Use K Means Clustering to cluster suburbs based on the similarity of venues
4. Summarise clusters, and determine the best cluster to open a bakery

Results

Sydney Suburb

Using The Suburbs in Sydney dataset, we can map the location of the suburbs in Sydney on the map.



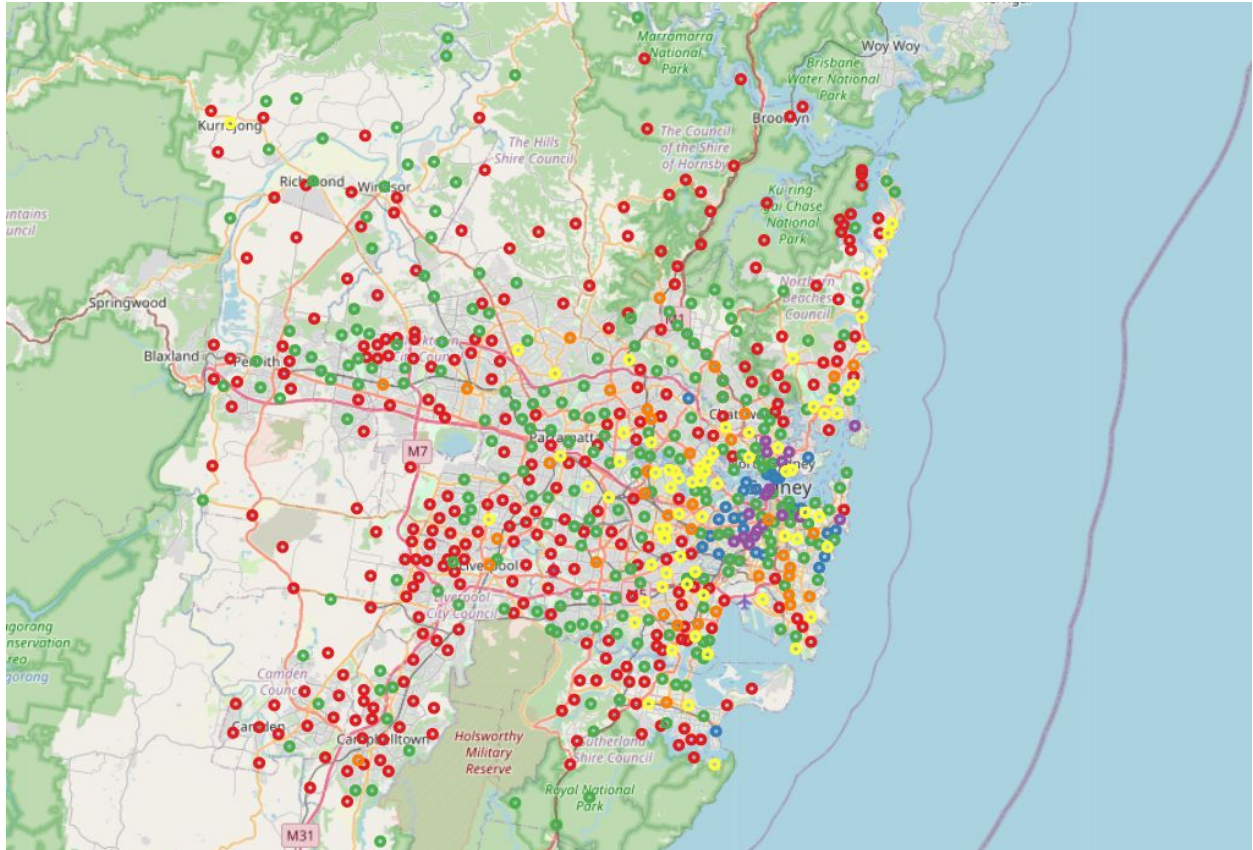
Venues

Using Foursquare Places API, we can obtain information on venues in the suburb

Suburb	Neighborhood	Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Abbotsbury		-33.869285	150.866703	Calmsley Hill City Farm	-33.868937	150.859516	Farm
Abbotsbury		-33.869285	150.866703	Abbotsbury Shops	-33.869554	150.865294	Convenience Store
Abbotsbury		-33.869285	150.866703	Stockdale Reserve	-33.871098	150.868060	Park
Abbotsbury		-33.869285	150.866703	Sydney United Sports Centre	-33.875559	150.874182	Athletics & Sports
Abbotsford		-33.850553	151.129759	The Cove Dining Co	-33.850950	151.129110	Café

Clustering

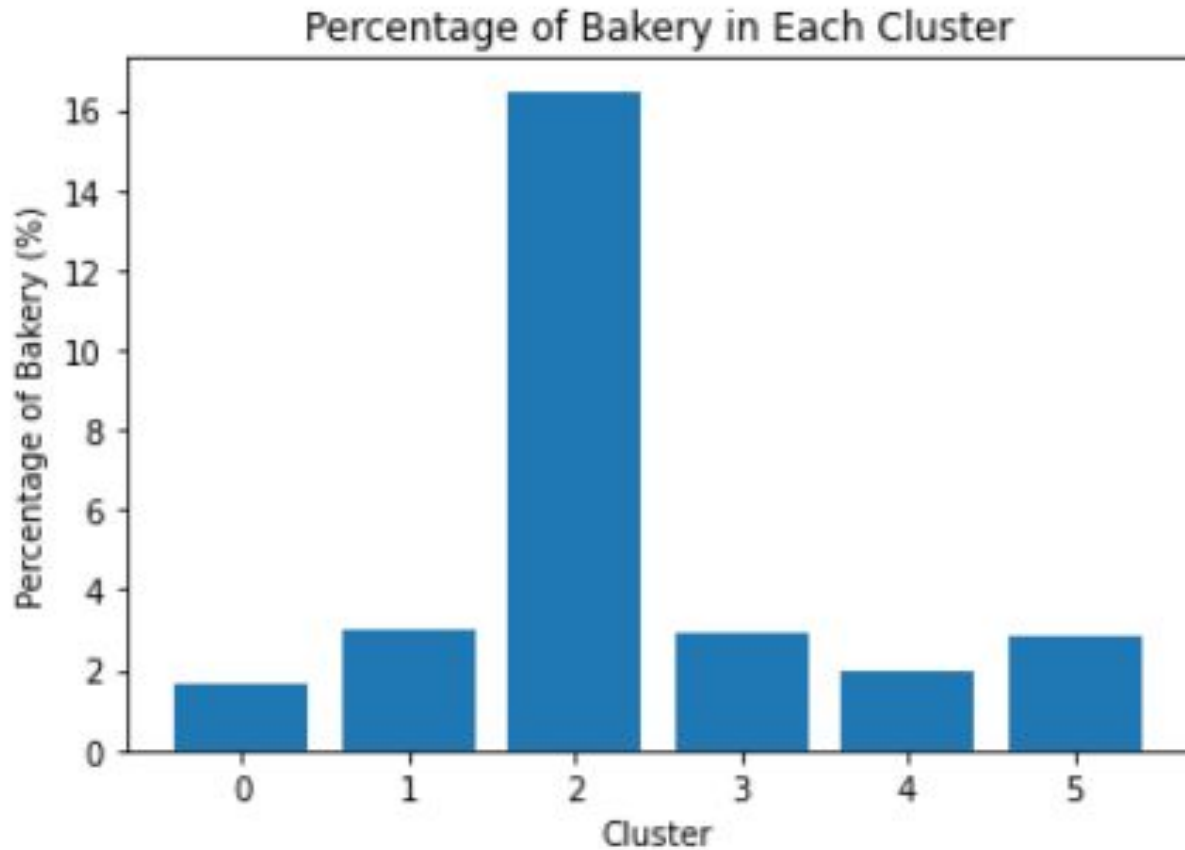
Create 6 clusters of suburbs based on venue similarities. This map shows the location of the clusters. Note that cluster 1 (Blue) and 3 (Purple) are located around the CBD area.



Cluster Top Venues

Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Fast Food Restaurant	Park	Café	Pizza Place	Grocery Store
1	Café	Park	Italian Restaurant	Bakery	Bar
2	Italian Restaurant	Thai Restaurant	Park	Bakery	Athletics & Sports
3	Café	Bar	Pub	Thai Restaurant	Japanese Restaurant
4	Café	Fast Food Restaurant	Chinese Restaurant	Supermarket	Coffee Shop
5	Café	Park	Pizza Place	Bakery	Beach
6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue	
Shopping Mall	Supermarket	Convenience Store	Train Station	Gym	
Pub	Australian Restaurant	Japanese Restaurant	Coffee Shop	Pizza Place	
Playground	Supermarket	Fast Food Restaurant	Shopping Mall	Lebanese Restaurant	
Coffee Shop	Italian Restaurant	Australian Restaurant	Ice Cream Shop	Bakery	
Thai Restaurant	Gym	Pizza Place	Japanese Restaurant	Vietnamese Restaurant	
Grocery Store	Supermarket	Thai Restaurant	Fast Food Restaurant	Italian Restaurant	

Bakeries In Each Cluster



Conclusion

Based on the results above, we can conclude:

- Cafe is the most popular venue in for clusters
- Based on the map computed, it seems location effects the venues in suburbs
- Cluster 0: has the lowest competition in bakery business compared to the other clusters. Retail businesses and the train station are among the most common in the area, while bakery is not in the top 10.
- Cluster 1: Bakery is the 4th most common venue in this area although the percentage of bakery is relatively low. Suburbs of this cluster are distributed around CBD.

- Cluster 2: Has the highest competition in bakery business. Bakery is the 4th most common venue in this cluster. Due to the competition, it might be unwise to open a new bakery in this area.
- Cluster 3: Similarly to Cluster 1, bakery is among the most common venues although the competition is relatively low. Most suburbs in this cluster are distributed around CBD
- Cluster 4: The majority of the most common venues in these clusters are food service and the competition for bakery is relatively low
- Cluster 5: Bakery is the 4th most common venue in this area although the percentage of bakery is relatively low. Suburbs of this cluster are distributed around beaches and parks.

In conclusion, I would recommend opening a bakery in suburbs in cluster 0