IBM Data Science Capstone Project

By Nicolas Dinh

Introduction

This project concludes of the IBM Data Science Certification Specialization. This goal of this last module is about to apply the knowledge and skills acquired during the course to a real-life problem. This project will be about solving a hypothetical business problem with real data (geolocation data from Foursquare using RESTful API calls as well as data scrapped directly from the web using libraries like BeautifulSoup) and analyzing it using Python in a Jupyter Notebook.

Similar to the approach that was taken during the course with New York and Toronto, the focus of this project will be about finding an optimal neighborhood (suburb) in the city of Melbourne in order to open a French restaurant.

Business Problem

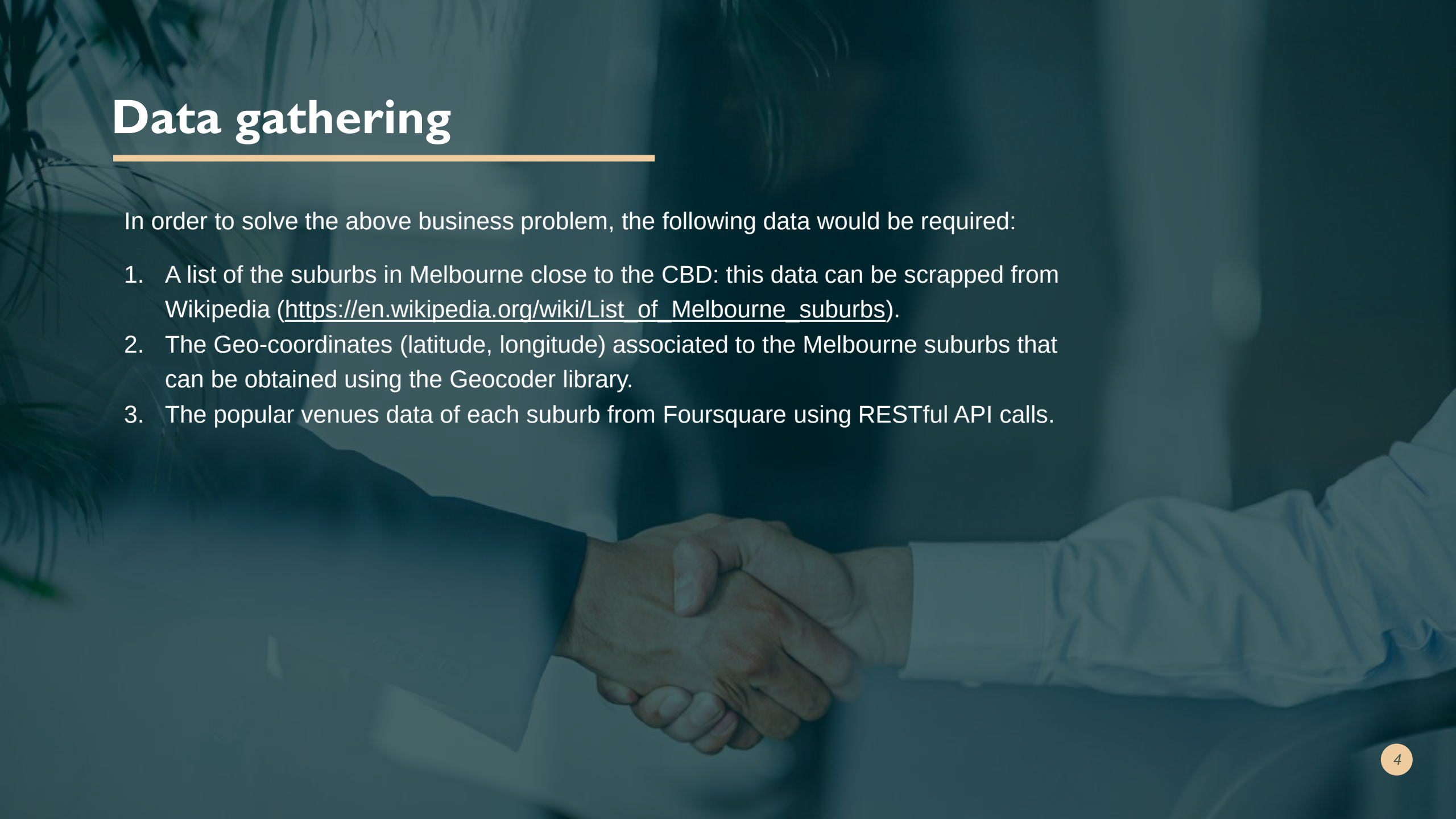
Melbourne is a city with a very multicultural population without over 200 nationalities and over 230 languages spoken. The city is the nest of a large amounts of communities and hosts numerous cultural events from music, international festivals to shows, expositions and galleries. And one of the big results of this multicultural aspect is the diversity of restaurants and food that the city has to offer.

My client, a French chef and entrepreneur asked me to take a data-driven approach to find a good place to open a French restaurant. Ideally, it would close to the city center and in a popular suburb outside of the CBD that already offer a large choice of restaurants, bars and cultural events/gathering.

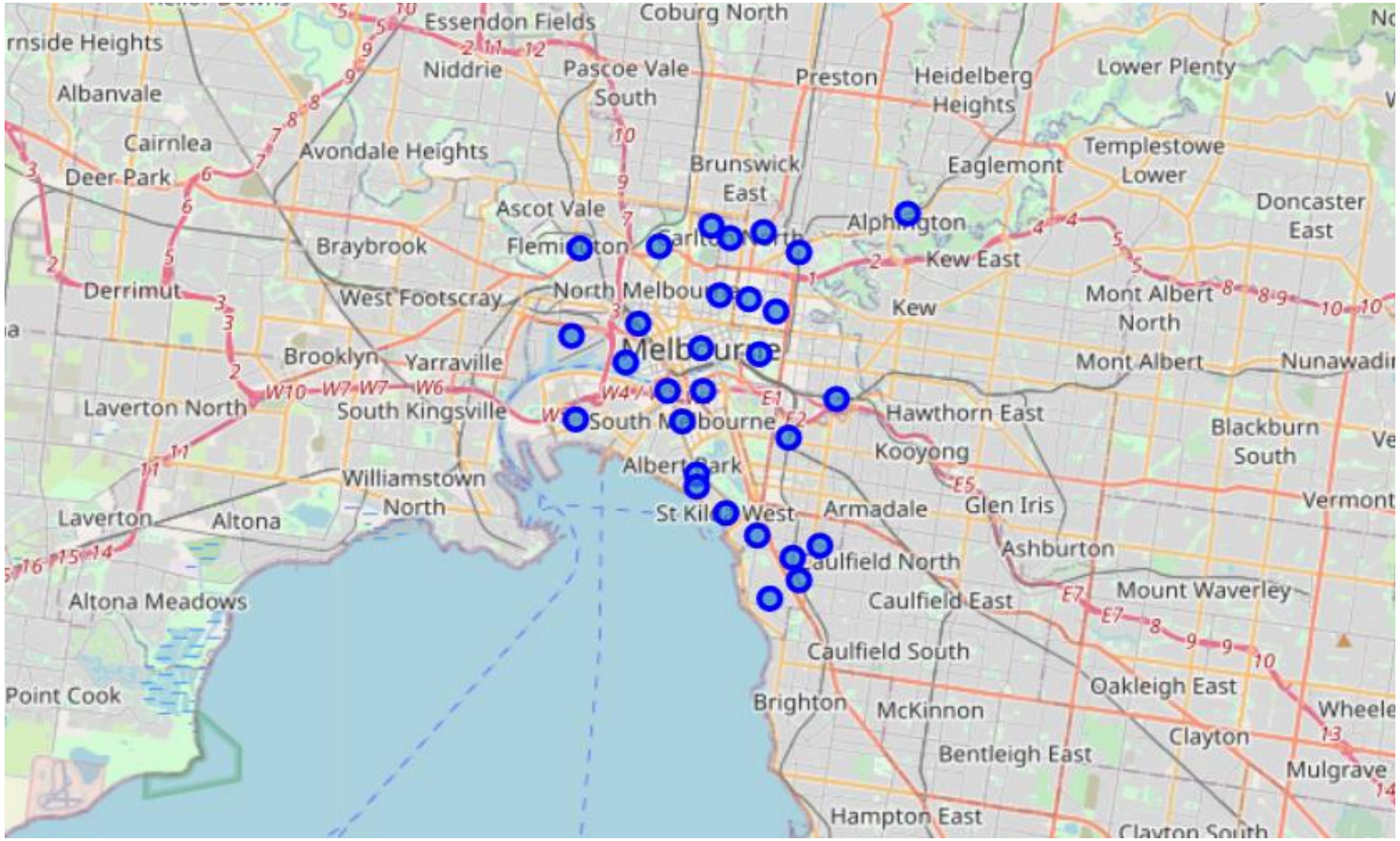
Data gathering

In order to solve the above business problem, the following data would be required:

1. A list of the suburbs in Melbourne close to the CBD: this data can be scrapped from Wikipedia (https://en.wikipedia.org/wiki/List_of_Melbourne_suburbs).
2. The Geo-coordinates (latitude, longitude) associated to the Melbourne suburbs that can be obtained using the Geocoder library.
3. The popular venues data of each suburb from Foursquare using RESTful API calls.



Visualization of the suburbs on a map

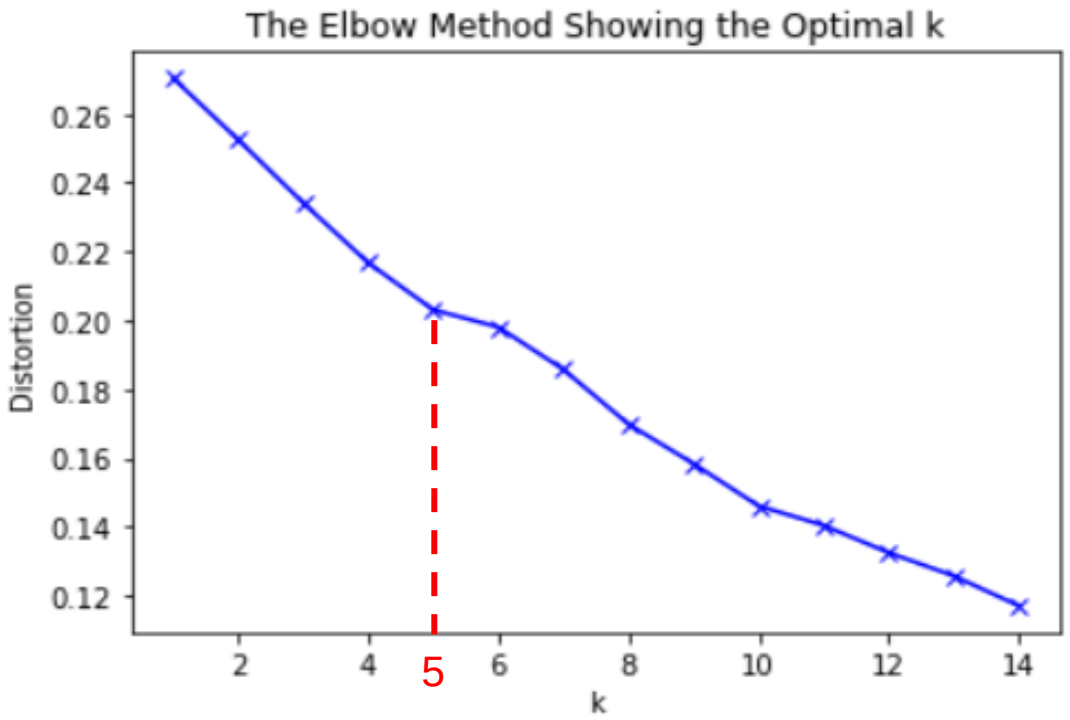


Most frequent venue category in each suburbs based on the Foursquare data

After exporting the list of most popular venues, we're looking at the frequency of occurrence of each category of venues as per below:

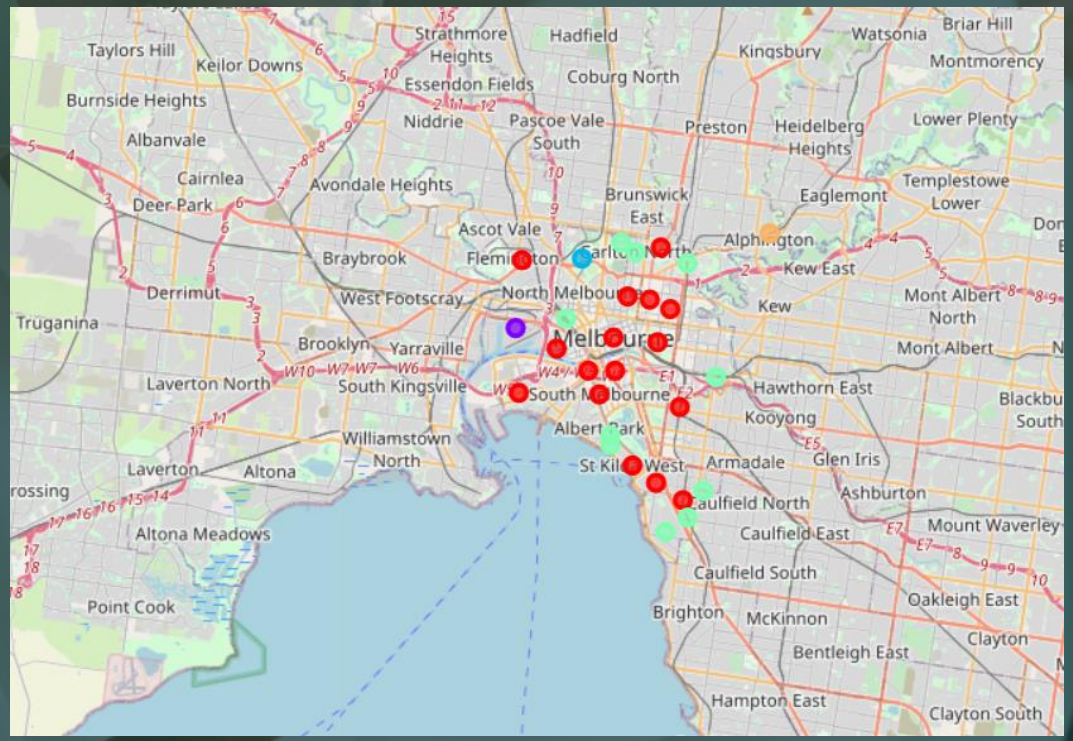
	Suburb	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Abbotsford	Café	Wine Shop	Italian Restaurant	Pizza Place	Park	Thai Restaurant	Zoo Exhibit	Flea Market	Fruit & Vegetable Store	Frozen Yogurt Shop
1	Albert Park	Café	Metro Station	Tennis Court	Golf Course	Seafood Restaurant	Food & Drink Shop	Athletics & Sports	Hotel	Indian Restaurant	Racetrack
2	Alphington	Liquor Store	Gym / Fitness Center	Convenience Store	Farmers Market	Fast Food Restaurant	Park	Train Station	Thai Restaurant	Flower Shop	Fruit & Vegetable Store
3	Balaclava	Café	Coffee Shop	Breakfast Spot	Bar	Pharmacy	Pizza Place	Vietnamese Restaurant	Tram Station	Salad Place	Japanese Restaurant
4	Burnley	Café	Pub	Furniture / Home Store	Breakfast Spot	Convenience Store	Park	Cocktail Bar	Shop & Service	Liquor Store	Food & Drink Shop

K-means clustering – choosing the optimal k



Model prediction and visualisation

	Postal Code	Municipality	Suburb	Latitude	Longitude	Cluster Labels	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	3053	City of Melbourne	Carlton	-37.800144	144.969268	0	Italian Restaurant	Café	Coffee Shop	Ice Cream Shop	Gourmet Shop	French Restaurant	Indie Theater	Park	Cheese Shop	Pub
1	3054	City of Melbourne	Carlton North	-37.784559	144.972855	3	Café	Bakery	Tram Station	Wine Bar	Flower Shop	Pub	Deli / Bodega	Italian Restaurant	Grocery Store	Liquor Store
2	3008	City of Melbourne	Docklands	-37.817870	144.938654	0	Italian Restaurant	Middle Eastern Restaurant	Café	Restaurant	Pier	Chinese Restaurant	Seafood Restaurant	Steakhouse	Sandwich Place	Tapas Restaurant
3	3002	City of Melbourne	East Melbourne	-37.815588	144.982555	0	Café	Cricket Ground	Hotel	Restaurant	Bar	Sushi Restaurant	Museum	Stadium	Coffee Shop	Tourist Information Center
4	3031	City of Melbourne	Flemington	-37.787277	144.923242	0	Hotel	Racecourse	Pizza Place	Tram Station	Park	Bowling Green	Pharmacy	Supermarket	Café	French Restaurant
5	3031	City of Melbourne	Kensington	-33.911719	151.219506	0	Pizza Place	Sushi Restaurant	Chinese Restaurant	Convenience Store	Kids Store	Gym / Fitness Center	Cosmetics Shop	Malay Restaurant	Liquor Store	Burger Joint
6	3000	City of Melbourne	Melbourne	-37.814218	144.963161	0	Coffee Shop	Bar	Café	Cocktail Bar	Dessert Shop	Shopping Mall	Clothing Store	Italian Restaurant	Cosmetics Shop	Candy Store
7	3051	City of Melbourne	North Melbourne	-37.807609	144.942351	3	Café	Sandwich Place	Theater	Zoo Exhibit	Flea Market	Furniture / Home Store	Fruit & Vegetable Store	Frozen Yogurt Shop	Fried Chicken Joint	French Restaurant
8	3052	City of Melbourne	Parkville	-37.787250	144.949596	2	Zoo Exhibit	Park	Hockey Arena	Gift Shop	Food & Drink Shop	Sculpture Garden	Basketball Court	Fast Food Restaurant	BBQ Joint	Sports Club
9	3207	City of Melbourne	Port Melbourne	-37.833361	144.921920	0	Climbing Gym	Café	Go Kart Track	Latin American Restaurant	Beach	Zoo Exhibit	Furniture / Home Store	Fruit & Vegetable Store	Frozen Yogurt Shop	Fried Chicken Joint



Recommendation

Most popular suburbs

Suburb	Venue
Melbourne	100
Fitzroy	100
South Yarra	91
South Melbourne	60
Collingwood	57
Balaclava	53
Southbank	47
Carlton	47
Docklands	44
Fitzroy North	39

Cluster 1 – Gastronomical center

	Municipality	Suburb	Cluster Labels	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	City of Melbourne	Carlton	0	Italian Restaurant	Café	Coffee Shop	Ice Cream Shop	Gourmet Shop	French Restaurant	Indie Theater	Park	Cheese Shop	Pub
2	City of Melbourne	Docklands	0	Italian Restaurant	Middle Eastern Restaurant	Café	Restaurant	Pier	Chinese Restaurant	Seafood Restaurant	Steakhouse	Sandwich Place	Tapas Restaurant
3	City of Melbourne	East Melbourne	0	Café	Cricket Ground	Hotel	Restaurant	Bar	Sushi Restaurant	Museum	Stadium	Coffee Shop	Tourist Information Center
4	City of Melbourne	Flemington	0	Hotel	Racecourse	Pizza Place	Tram Station	Park	Bowling Green	Pharmacy	Supermarket	Café	French Restaurant
5	City of Melbourne	Kensington	0	Pizza Place	Sushi Restaurant	Chinese Restaurant	Convenience Store	Kids Store	Gym / Fitness Center	Cosmetics Shop	Malay Restaurant	Liquor Store	Burger Joint
6	City of Melbourne	Melbourne	0	Coffee Shop	Bar	Café	Cocktail Bar	Dessert Shop	Shopping Mall	Clothing Store	Italian Restaurant	Cosmetics Shop	Candy Store
9	City of Melbourne	Port Melbourne	0	Climbing Gym	Café	Go Kart Track	Latin American Restaurant	Beach	Zoo Exhibit	Furniture / Home Store	Fruit & Vegetable Store	Frozen Yogurt Shop	Fried Chicken Joint
10	City of Melbourne	Southbank	0	Café	Hotel	Grocery Store	Bar	Performing Arts Venue	Italian Restaurant	Restaurant	Coffee Shop	Bakery	Australian Restaurant
11	City of Melbourne	South Wharf	0	Hotel	Bar	Clothing Store	Sporting Goods Shop	Australian Restaurant	Multiplex	Restaurant	Plaza	Seafood Restaurant	Shoe Store
12	City of Melbourne	South Yarra	0	Café	Italian Restaurant	Hotel	Japanese Restaurant	Grocery Store	Bakery	Convenience Store	Coffee Shop	Pizza Place	Dessert Shop
15	City of Port Phillip	Balaclava	0	Café	Coffee Shop	Breakfast Spot	Bar	Pharmacy	Pizza Place	Vietnamese Restaurant	Tram Station	Salad Place	Japanese Restaurant
19	City of Port Phillip	St Kilda	0	Café	Fast Food Restaurant	Convenience Store	Japanese Restaurant	Pub	Tram Station	Australian Restaurant	Thai Restaurant	Pizza Place	Theater
21	City of Port Phillip	St Kilda West	0	Café	Tram Station	Italian Restaurant	Hotel Bar	Australian Restaurant	Garden	Music Venue	Beach	Restaurant	Juice Bar
22	City of Port Phillip	South Melbourne	0	Café	Bar	Coffee Shop	Gastropub	Mexican Restaurant	Bakery	Spa	Fish & Chips Shop	Breakfast Spot	Pub
27	City of Yarra	Collingwood	0	Café	Bar	Japanese Restaurant	Gay Bar	Cocktail Bar	Bakery	Kebab Restaurant	Pub	Coffee Shop	Adult Boutique
28	City of Yarra	Cremorne	0	Café	Bar	Park	Shopping Mall	Sushi Restaurant	Korean Restaurant	Grocery Store	Cocktail Bar	Fast Food Restaurant	Movie Theater
30	City of Yarra	Fitzroy	0	Café	Bar	Cocktail Bar	Vietnamese Restaurant	Pub	Bakery	Bookstore	Japanese Restaurant	Wine Bar	Vegetarian / Vegan Restaurant
31	City of Yarra	Fitzroy North	0	Coffee Shop	Café	Tram Station	Pub	Ice Cream Shop	Bakery	Italian Restaurant	Park	Veterinarian	Beer Garden

Based on the above cluster, we can definitely recommend the 1st and 4th cluster to open a restaurant. However, based on a previous analysis we could see that Melbourne, Fitzroy and South Yarra are the suburbs with the highest density of restaurants, cafes and bars in a 500m radius. These 3 suburbs all belong to the 1st cluster. As the client wants to avoid opening a restaurant directly in the CBD but still wants proximity to it, **the best suburbs to open a French restaurant would be Fitzroy or South Yarra**.

Conclusion

In this project, we used different concepts and python libraries to come up with an answer to a real-life like business problem and use data from renown sources like Wikipedia and Foursquare. We've been able to use data science concepts that we learn along the specialization and apply it for this capstone project:

- Data collection (using web-scraping techniques and RESTful API calls)
- Data cleaning and preprocessing using pandas
- Geo-coordinates data Visualization using Folium
- Unsupervised machine learning to perform clustering using scikit-learn

All of these contributed to provide a recommendation to the business problem.

Thank you for reading through all my project