

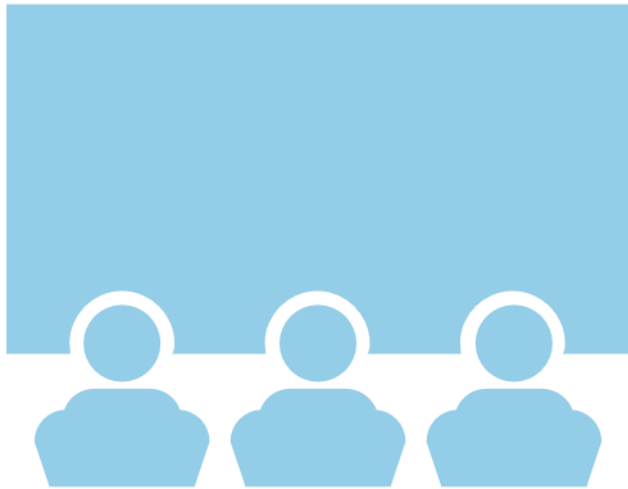
Data Science Capstone project

Analyzing the Neighbourhoods in Mumbai to Start a New Restaurant

ERIC KWASI GADOSEY

6TH SEPTEMBER, 2021

Outline



- Executive Summary
- Introduction
- Methodology
- Results
- Conclusion
- Appendix

Executive Summary



- Summary of methodologies

The main objective of this project is to carefully analyze appropriate data and find recommendations for the stakeholders.

- Summary of all results

Using data science methodology and machine learning techniques like clustering, for this project

Introduction



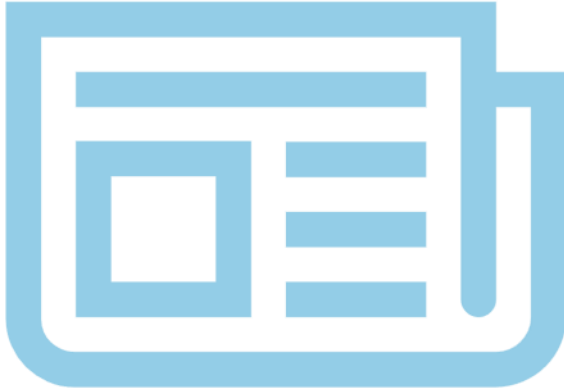
- **Project background and context**

Mumbai is India's largest city by population and also is the financial and commercial capital of the country as it generates 6.16% of the total Gross Domestic Product (GDP). It lies on the west coast of India and attracts heavy tourism from all over the globe every year. It is one of the major hubs of the world and is extremely diverse with people from various ethnicities residing there. The multi-cultural nature of the city of Mumbai has brought along with it numerous cuisines from all over the world. Mumbai has a variety of street food, including the Vada pav, Pav bhaji, Chaat, Idli, Dosa and the Mumbai sandwich. Thus, this project aims to study the neighbourhoods in Mumbai to determine possible locations for starting a restaurant.

- **Problems you want to find answers**

Data that might contribute to determining if any neighbourhoods in Mumbai are most possible and efficient to start up a restaurant. This project aims to predict whether and how the new restaurant should be start up to accommodate their cultural settings.

Methodology



- **Data collection methodology:**
- Neighbourhood data of Mumbai
- Geographical coordinates of Mumbai and all neighbourhoods in Mumbai
- Venue data for neighbourhoods in Mumbai

Methodology

This section provides details of the methodology used in the project.

Data collection

- Describe how data sets were collected.

The data of the neighbourhoods in Mumbai was scraped from https://en.wikipedia.org/wiki/List_of_neighborhoods_in_Mumbai.

- You need to present your data collection process use key phrases and flowcharts

The data is read into a pandas data frame using the `read_html()` method. The main reason for doing so is that the Wikipedia page provides a comprehensive and detailed table of the data which can easily be scraped using the `read_html()` method of pandas.

The data of the neighbourhoods in Mumbai was scraped from https://en.wikipedia.org/wiki/List_of_neighborhoods_in_Mumbai ■

Data collection – web scraping

Figure 1: Top 10 rows of Mumbai neighbourhoods data scraped from Wikipedia.

	Neighborhood	Location	Latitude	Longitude
0	Amboli	Andheri, Western Suburbs	19.129300	72.843400
1	Chakala, Andheri	Western Suburbs	19.111388	72.860833
2	D.N. Nagar	Andheri, Western Suburbs	19.124085	72.831373
3	Four Bungalows	Andheri, Western Suburbs	19.124714	72.827210
4	Lokhandwala	Andheri, Western Suburbs	19.130815	72.829270
5	Marol	Andheri, Western Suburbs	19.119219	72.882743
6	Sahar	Andheri, Western Suburbs	19.098889	72.867222
7	Seven Bungalows	Andheri, Western Suburbs	19.129052	72.817018
8	Versova	Andheri, Western Suburbs	19.120000	72.820000
9	Mira Road	Mira-Bhayandar, Western Suburbs	19.284167	72.871111

The code for getting the geographical coordinates of Mumbai is shown in Figure 2.

```
address = 'Mumbai, IN'
geolocator = Nominatim()
location = geolocator.geocode(address)
latitude = location.latitude
longitude = location.longitude
print('The geograpical coordinates of Mumbai are {}, {}'.format(latitude, longitude))
```

The geograpical coordinates of Mumbai are 19.0759899, 72.8773928.

Figure 2: Obtaining geographical coordinates of Mumbai.

Using folium, a map was plotted to show how the different neighbourhoods are spread all across Mumbai. This is shown in Figure 7.

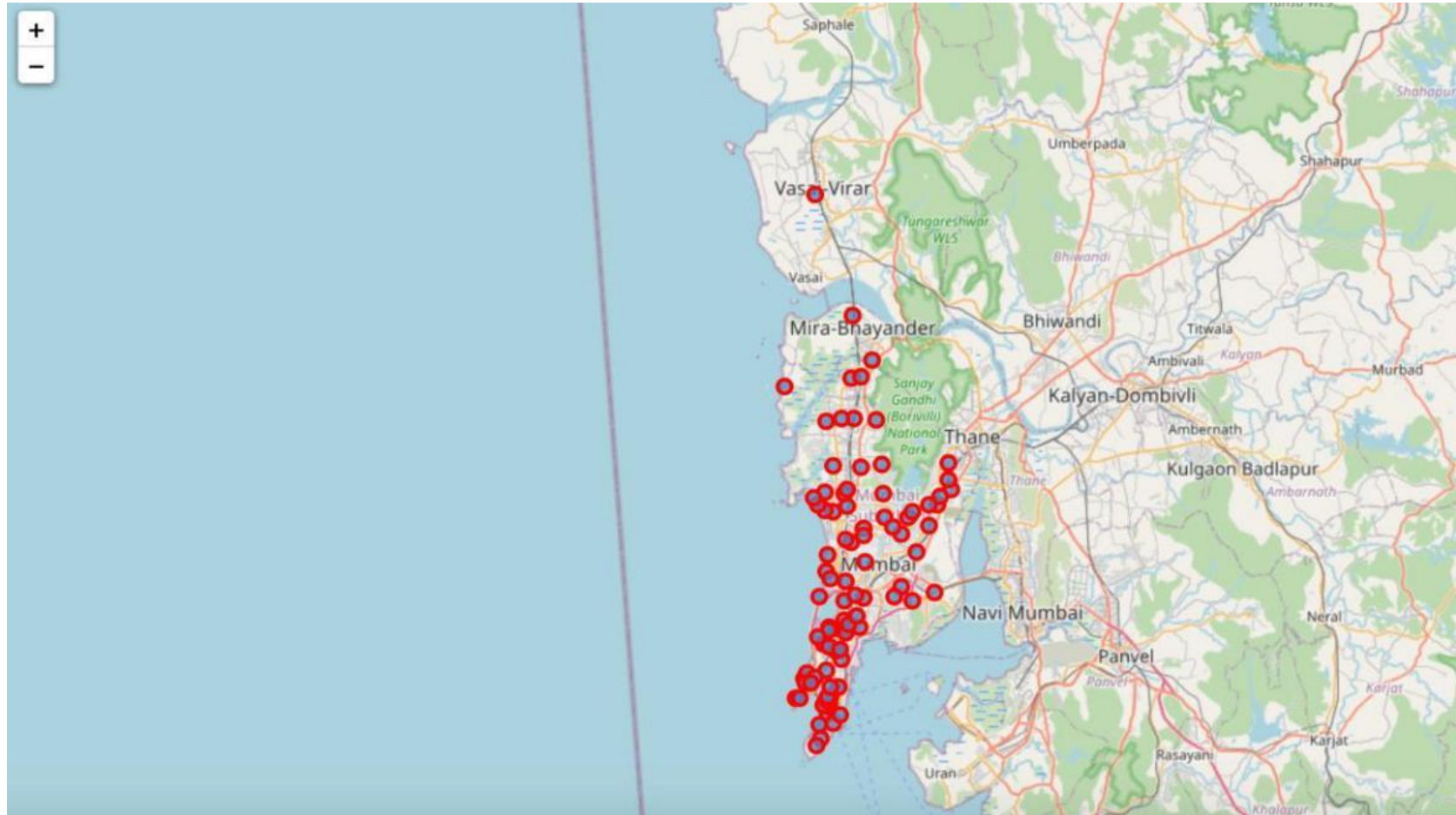


Figure 7: Depicting the neighbourhood spread across Mumbai.

Predictive Modeling

K-means unsupervised learning technique was used to cluster the neighbourhoods based on the category of venues near the neighbourhoods.

One important aspect of the k-means model is to determine the number of clusters to use in model development.

This was determined by the Silhouette score which was calculated for a range of clusters from 2 to 15.

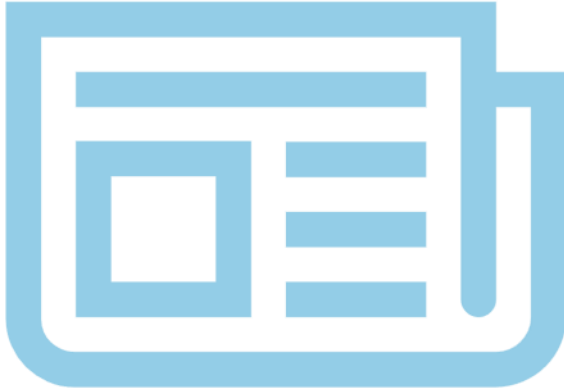
The resulting number of clusters and their respective Silhouette scores are shown in Figure 10.

Predictive analysis (Clustering)



Figure 10: Silhouette scores for different clusters

Results



The clustering model then clusters the neighbourhoods in Mumbai and provides a label for each neighbourhood that is representative of the cluster it belongs to.

Location, Latitude, and Longitude columns to provide a complete summary of the clustering. The top 10 rows are shown in Figure 11.

	Neighborhood	Location	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	Amboli	Western Suburbs	19.1293	72.8464	1	Indian Restaurant	Coffee Shop	Bakery	Bar	Asian Restaurant	Pizza Place	Sandwich Place	Bowling Alley	Bus Station	Bike Rental / Bike Share
1	Chakala, Andheri	Western Suburbs	19.1084	72.8623	1	Hotel	Indian Restaurant	Café	Fast Food Restaurant	Pizza Place	Asian Restaurant	Hotel Bar	Vegetarian / Vegan Restaurant	Restaurant	Gym
2	D.N. Nagar	Western Suburbs	19.1241	72.8325	0	Bar	Indian Restaurant	Pub	Gym / Fitness Center	Pizza Place	Lounge	Coffee Shop	Vegetarian / Vegan Restaurant	Snack Place	Gym
3	Four Bungalows	Western Suburbs	19.1263	72.8243	0	Pub	Café	Indian Restaurant	Gym / Fitness Center	Chinese Restaurant	Bar	Seafood Restaurant	Lounge	Vegetarian / Vegan Restaurant	Coffee Shop
4	Lokhandwala	Western Suburbs	19.1432	72.8249	0	Indian Restaurant	Chinese Restaurant	Café	Pub	Bakery	Bar	Italian Restaurant	Gym / Fitness Center	Coffee Shop	Asian Restaurant
5	Marol	Western Suburbs	19.1192	72.8827	1	Indian Restaurant	Hotel	Diner	Bakery	Dance Studio	Ice Cream Shop	Chinese Restaurant	Fast Food Restaurant	Restaurant	Lounge
6	Sahar	Western Suburbs	19.1027	72.8626	0	Hotel	Café	Indian Restaurant	Lounge	Gym	Asian Restaurant	Pizza Place	Seafood Restaurant	Restaurant	Falafel Restaurant
7	Seven Bungalows	Western Suburbs	19.1315	72.817	0	Café	Pub	Seafood Restaurant	Chinese Restaurant	Pizza Place	Coffee Shop	Bar	Ice Cream Shop	Asian Restaurant	Bistro
8	Versova	Western Suburbs	19.1377	72.8135	0	Café	Ice Cream Shop	Beach	Pizza Place	Coffee Shop	Chinese Restaurant	Salon / Barbershop	Frozen Yogurt Shop	Bistro	Sandwich Place
9	Mira Road	Western Suburbs	19.2657	72.8711	1	Indian Restaurant	Convenience Store	Coffee Shop	Mexican Restaurant	Fast Food Restaurant	Food Truck	Motorcycle Shop	Movie Theater	Basketball Court	Bar

Figure 11: Clustering neighbourhoods in Mumbai.

Furthermore, neighbourhoods in each cluster can be extracted using cluster labels and thus the details of specific clusters can be seen. This is done below for all clusters with only 10 rows for clusters that contain a high number of neighbourhoods.

	Neighborhood	Location	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
1	Chakala, Andheri	Western Suburbs	Hotel	Indian Restaurant	Café	Hotel Bar	Asian Restaurant	Pizza Place	Vegetarian / Vegan Restaurant	Restaurant	Burger Joint	Multiplex
6	Sahar	Western Suburbs	Hotel	Indian Restaurant	Restaurant	Gym	Asian Restaurant	Bar	Coffee Shop	Café	Italian Restaurant	Pub
27	Khar Danda	Western Suburbs	Hotel	Clothing Store	Park	Coffee Shop	Dessert Shop	Bookstore	Bistro	French Restaurant	Boutique	Pool
40	Kanjurmarg	Eastern Suburbs	Train Station	Gym	Hotel	Gift Shop	Chinese Restaurant	French Restaurant	Asian Restaurant	Multiplex	Donut Shop	Electronics Store
70	Malabar Hill	South Mumbai	Gym	Hotel	Park	Convenience Store	Lighthouse	Coffee Shop	Dessert Shop	Indian Restaurant	Cupcake Shop	Cosmetics Shop
77	Walkeshwar	South Mumbai	Gym	Park	Hotel	Convenience Store	Food & Drink Shop	Food Truck	Lighthouse	Restaurant	Dessert Shop	Coffee Shop

Figure 12: Cluster 1.

	Neighborhood	Location	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	Amboli	Western Suburbs	Indian Restaurant	Pizza Place	Sandwich Place	Bar	Bakery	Athletics & Sports	Metro Station	Burger Joint	Bike Rental / Bike Share	Mughlai Restaurant
2	D.N. Nagar	Western Suburbs	Pub	Bar	Indian Restaurant	Pizza Place	Gym / Fitness Center	Women's Store	Vegetarian / Vegan Restaurant	Juice Bar	Snack Place	Hotel Bar
3	Four Bungalows	Western Suburbs	Pub	Lounge	Chinese Restaurant	Indian Restaurant	Seafood Restaurant	Vegetarian / Vegan Restaurant	Pizza Place	Café	Coffee Shop	Ice Cream Shop
4	Lokhandwala	Western Suburbs	Indian Restaurant	Café	Coffee Shop	Chinese Restaurant	Bar	Pub	Italian Restaurant	Fast Food Restaurant	Asian Restaurant	Seafood Restaurant
7	Seven Bungalows	Western Suburbs	Café	Pub	Bar	Chinese Restaurant	Indian Restaurant	Seafood Restaurant	Pizza Place	Ice Cream Shop	Vegetarian / Vegan Restaurant	Coffee Shop
...
83	Ballard Estate	South Mumbai	Indian Restaurant	Café	Coffee Shop	Seafood Restaurant	Bar	Parsi Restaurant	Lounge	Train Station	Irani Cafe	Dessert Shop
88	Parel	South Mumbai	Coffee Shop	Playground	Indian Restaurant	Chinese Restaurant	Plaza	Restaurant	Rest Area	Maharashtrian Restaurant	Bar	Vegetarian / Vegan Restaurant
89	Gowalia Tank	South Mumbai	Indian Restaurant	Fast Food Restaurant	Coffee Shop	Bakery	Electronics Store	Café	Sandwich Place	Vegetarian / Vegan Restaurant	Snack Place	Ice Cream Shop
90	Dava Bazaar	South Mumbai	Train Station	Indian Restaurant	Fish Market	Hotel	French Restaurant	Fast Food Restaurant	Café	Coffee Shop	Clothing Store	Asian Restaurant
91	Dharavi	Mumbai	Indian Restaurant	Paper / Office Supplies Store	Café	Fast Food Restaurant	Shoe Store	Seafood Restaurant	Sandwich Place	Luggage Store	Diner	Bus Station

Figure 13: Cluster 2.

	Neighborhood	Location	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
92	Thane	Mumbai	Platform	Pizza Place	Performing Arts Venue	Dessert Shop	Fish & Chips Shop	Field	Fast Food Restaurant	Falafel Restaurant	Event Space	Electronics Store

Figure 14: Cluster 3.

	Neighborhood	Location	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
5	Marol	Western Suburbs	Indian Restaurant	Chinese Restaurant	Hotel	Coffee Shop	Bakery	Restaurant	Lounge	Department Store	Dance Studio	Diner
9	Mira Road	Western Suburbs	Indian Restaurant	Mexican Restaurant	Bar	Coffee Shop	Fast Food Restaurant	Movie Theater	Basketball Court	Convenience Store	Gym / Fitness Center	Dhaba
11	Uttan	Western Suburbs	Convenience Store	Indian Restaurant	Whisky Bar	Restaurant	Zoo	Dhaba	Fish Market	Fish & Chips Shop	Field	Fast Food Restaurant
20	Jogeshwari West	Western Suburbs	Indian Restaurant	Restaurant	Gym	Café	Business Service	Bus Station	Men's Store	Mughlai Restaurant	Ice Cream Shop	Chinese Restaurant
36	Bhandup	Eastern Suburbs	Train Station	Indian Restaurant	Fast Food Restaurant	Asian Restaurant	Zoo	Dim Sum Restaurant	Flea Market	Fish Market	Fish & Chips Shop	Field
37	Amrut Nagar	Eastern Suburbs	Indian Restaurant	Ice Cream Shop	Pharmacy	Train Station	Café	Multiplex	Bus Station	Mughlai Restaurant	Restaurant	Plaza
38	Pant Nagar	Eastern Suburbs	Indian Restaurant	Coffee Shop	Shopping Mall	Café	Vegetarian / Vegan Restaurant	Bakery	Platform	Fast Food Restaurant	Bus Station	Diner
43	Chandivali	Eastern Suburbs	Indian Restaurant	Fast Food Restaurant	Restaurant	Seafood Restaurant	Brewery	Café	Food	Lounge	Chinese Restaurant	Pub
46	VidyaVihar	Eastern Suburbs	Indian Restaurant	Fast Food Restaurant	Juice Bar	Playground	Chinese Restaurant	Dessert Shop	Ice Cream Shop	Vegetarian / Vegan Restaurant	Train Station	Bar
48	Chembur	Harbour Suburbs	Indian Restaurant	Pizza Place	Café	Fast Food Restaurant	Vegetarian / Vegan Restaurant	Dessert Shop	Asian Restaurant	Bar	Seafood Restaurant	Garden
51	Mahul	Harbour Suburbs	Indian Restaurant	Fast Food Restaurant	Smoke Shop	Hotel	Pizza Place	Cosmetics Shop	Ice Cream Shop	Vegetarian / Vegan Restaurant	Pool	Bar
54	Bhuleshwar	South Mumbai	Indian Restaurant	Bakery	Chinese Restaurant	Market	Café	Ice Cream Shop	Dessert Shop	Multiplex	Fast Food Restaurant	Rest Area
57	Cavel	South Mumbai	Indian Restaurant	Bakery	Café	Chinese Restaurant	Market	Restaurant	Ice Cream Shop	Fast Food Restaurant	Multiplex	Arcade
63	Dhobitalao	South Mumbai	Indian Restaurant	Coffee Shop	Café	Bar	Multiplex	Flea Market	Chinese Restaurant	Train Station	Bakery	Market
64	Dongri	South Mumbai	Indian Restaurant	Dessert Shop	Chinese Restaurant	Convenience Store	Garden	Restaurant	Market	Café	BBQ Joint	Boat or Ferry
68	Mahalaxmi	South Mumbai	Indian Restaurant	Café	Chinese Restaurant	Multiplex	Bakery	Fast Food Restaurant	Ice Cream Shop	Bar	Snack Place	Jewelry Store
69	Mahim	South Mumbai	Indian Restaurant	Fast Food Restaurant	Bakery	Chinese Restaurant	Dessert Shop	Café	Clothing Store	Breakfast Spot	Coffee Shop	Asian Restaurant
71	Marine Drive	South Mumbai	Indian Restaurant	Bakery	Multiplex	Ice Cream Shop	Chinese Restaurant	Fast Food Restaurant	Café	Train Station	Cheese Shop	Market
78	Worli	South Mumbai	Indian Restaurant	Scenic Lookout	Playground	Café	Garden	Diner	Thai Restaurant	Sandwich Place	Food	Gourmet Shop
84	Chira Bazaar	South Mumbai	Indian Restaurant	Café	Bakery	Multiplex	Ice Cream Shop	Chinese Restaurant	Fast Food Restaurant	Gym	Juice Bar	Flea Market
85	Fanas Wadi	South Mumbai	Indian Restaurant	Café	Fast Food Restaurant	Chinese Restaurant	Market	Ice Cream Shop	Juice Bar	Dessert Shop	Train Station	Coffee Shop
86	Chor Bazaar	South Mumbai	Indian Restaurant	Dessert Shop	Middle Eastern Restaurant	Arts & Crafts Store	Rest Area	Café	BBQ Joint	Market	Chinese Restaurant	Fast Food Restaurant
87	Matunga	South Mumbai	Indian Restaurant	Fast Food Restaurant	Café	Snack Place	Ice Cream Shop	Gym	Miscellaneous Shop	Chinese Restaurant	Vegetarian / Vegan Restaurant	Train Station

Figure 15: Cluster 4.

	Neighborhood	Location	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
59	Cotton Green	South Mumbai	Plaza	Pizza Place	Fast Food Restaurant	Train Station	Bakery	Dhaba	Fish Market	Fish & Chips Shop	Field	Falafel Restaurant

Figure 16: Cluster 5.

Based on the clusters shown above, the neighbourhoods can once again be plotted on a map of Mumbai, however, this time with different colour markers to distinguish between different clusters.

CONCLUSION



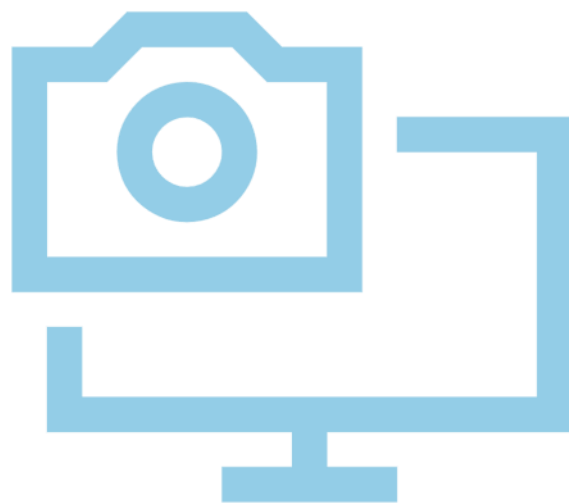
In this project, the neighbourhoods in Mumbai, India have been successfully analyzed for determining which would be the best neighbourhoods for opening a new restaurant.

Based on the analysis carried out, neighbourhoods in cluster 1 are recommended as locations for the new restaurant.

The stakeholders and investors can further tune this by considering various other factors like transport, legal requirements, and costs associated.

These were out of the scope for this project and thus were not considered.

APPENDIX



References

1. [*"Mumbai Suburban"* \(PDF\)](#). [*National Informatics Centre* \(Maharashtra State Centre\)](#). Archived from [*the original* \(PDF\)](#) on 13 May 2012. Retrieved 17 July 2013.
2. [*"Mumbai Urban Infrastructure Project"*](#). [*Mumbai Metropolitan Region Development Authority* \(MMRDA\)](#). Archived from the original on 26 February 2009. Retrieved 17 July 2013.
3. [*"Area and Density – Metropolitan Cities"* \(PDF\)](#). Ministry of Urban Development ([*Government of India*](#)). p. 33. Archived from [*the original* \(PDF\)](#) on 29 April 2009. Retrieved 17 July 2013.
4. [*Mumbai Plan*](#), 1.2 Area and Divisions
5. [*Greater Bombay District Gazetteer 1960*](#), p. 2
6. [*Mumbai Plan*](#), 1.1 Location
7. [*Bapat 2005*](#), pp. 111–112