The Battle of Neighborhoods:

Ideal location for Indian Fast food Joint

Vaibhavi Sonetha

July, 2019

Introduction

- Opening a successful restaurant is all about location.
- The right location is determined by a combination of restaurant concept and ideal customer.
- Defining the restaurant type and identifying the target demographics helps in selecting the correct restaurant location which sets the business up for the success.
- A young restaurateur has chain of fast food joints across India and wishes to start one at New York as it is a ‘City of dreams’.
- Generally, fast food is liked by age group of 15-35 years. It is a spontaneous outings that are often connected to another activity like shopping or an evening out with friends. Lower-income neighborhoods with high-population density and high foot traffic are key considerations for such joints.
- Data science with its established and effective analysis strategies can help the restaurateur to identify the ideal location for the new venture.

Problem Description

- Indian cuisine is elaborate and has varied dishes with Fast food being the most popular one amongst the young generation.
- More than 50% Indians residing in New York are concentrated at Queens Borough. Also, the Indian restaurants present at Queens mainly cater to fine dining. Hence, venturing out a Fast food joint will surely tap the market early.
- The restaurateur needs help in identifying the right location at Queens which can be done using the data science approach.

Problem Statement

- To identify the ideal location for opening Fast food joint at Queens Borough using principles and techniques of data science.

Data Sources

- New York data: https://cocl.us/new_york_dataset
- Data: 5 Boroughs of New York, 306 neighborhoods, geographical coordinates
- Queens Borough: 81 neighborhoods
- FOURSQUARE API : Venues in 500 meter radius of each neighborhood of Queens
- Dataset from FOURSQUARE: 2083 venues
- Unique categories of venues: ‘groupby’ function of Pandas

Feature Selection

1. No Indian restaurant in the neighborhood
2. Presence of beach
3. Presence of Bowling Alley
4. Presence of Flea market
5. Presence of Metro station
6. Presence of shopping mall
7. Presence of multiplex
8. Presence of park
9. Presence of scenic lookout
10. Movie theatre

Methodology

Step 1: Collection of New York data

Step 2: Geographical coordinates of Queens Borough and its neighborhoods

Step 3: Querying FOURSQUARE API to obtains venues in 500 meter radius of each neighborhood of Queens

Step 4: Group data by neighborhood and calculate total venues for each neighborhood

Step 5: Remove neighborhoods with Indian restaurants

Step 6: Feature selection, ‘Beach’, ‘Bowling Alley’, ‘Flea Market’, ‘Metro station’, ‘Shopping Mall’, ‘Multiplex, ‘Park’, ‘Scenic Lookout’ and ‘Movie Theater’

Step 7: k-means clustering to obtain neighborhoods with the maximum number of venues and maximum unique venues

Results

- New York data: 2090 rows indicating 5 boroughs, 306 neighborhoods

	Borough	Neighborhood	Latitude	Longitude
0	Bronx	Wakefield	40.894705	-73.847201
1	Bronx	Co-op City	40.874294	-73.829939
2	Bronx	Eastchester	40.887556	-73.827806
3	Bronx	Fieldston	40.895437	-73.905643
4	Bronx	Riverdale	40.890834	-73.912585

Results

- Data of Queens borough: 81 rows indicating 81 neighborhoods

	Borough	Neighborhood	Latitude	Longitude
0	Queens	Astoria	40.768509	-73.915654
1	Queens	Woodside	40.746349	-73.901842
2	Queens	Jackson Heights	40.751981	-73.882821
3	Queens	Elmhurst	40.744049	-73.881656
4	Queens	Howard Beach	40.654225	-73.838138

Results

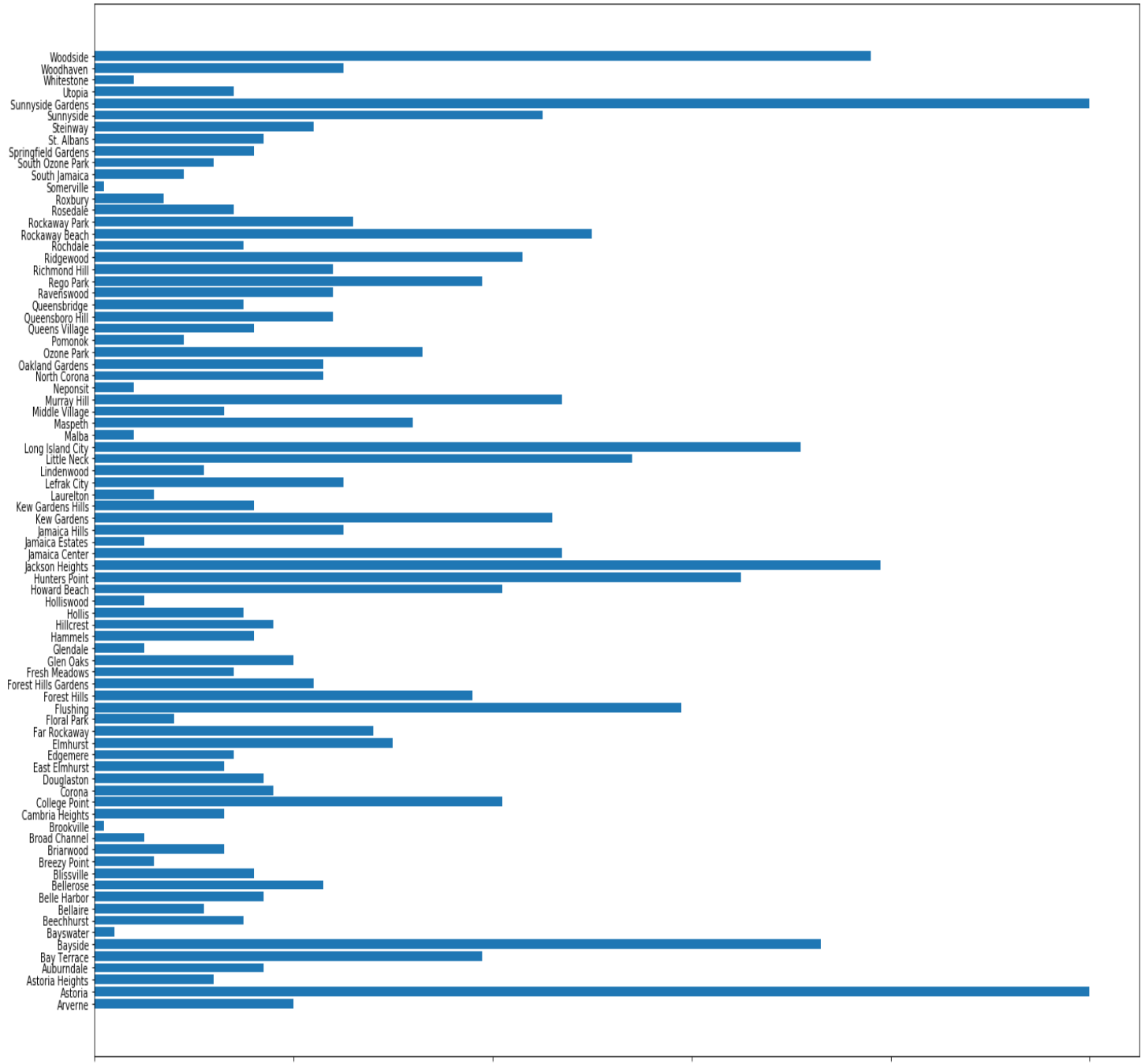
- Querying Foursquare API for each of 81 neighborhoods: 2083 rows indicating that 2083 venues

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Astoria	40.768509	-73.915654	Favela Grill	40.767348	-73.917897	Brazilian Restaurant
1	Astoria	40.768509	-73.915654	Orange Blossom	40.769856	-73.917012	Gourmet Shop
2	Astoria	40.768509	-73.915654	Simply Fit Astoria	40.769114	-73.912403	Gym
3	Astoria	40.768509	-73.915654	CrossFit Queens	40.769404	-73.918977	Gym
4	Astoria	40.768509	-73.915654	Titan Foods Inc.	40.769198	-73.919253	Gourmet Shop

- Grouping of venues by categories: ‘groupby’ function, 272 unique categories of venues

Results

- Bar plot of all neighborhoods vs unique categories
- Boroughs with maximum number of venues: ‘Astoria’, ‘Sunnyside Gardens’



Results

- Neighborhoods without Indian restaurant: 68 neighborhoods
- ‘Astoria’ and ‘Sunnyside Gardens’ has Indian restaurant and hence removed
- Featured venues: ‘Beach’, ‘Bowling Alley’, ‘Flea Market’, ‘Metro station’, ‘Shopping Mall’, ‘Multiplex, ‘Park’, ‘Scenic Lookout’, ‘Movie Theater’
- Neighborhoods with at least one featured venue: 36
- Featured venues: 9

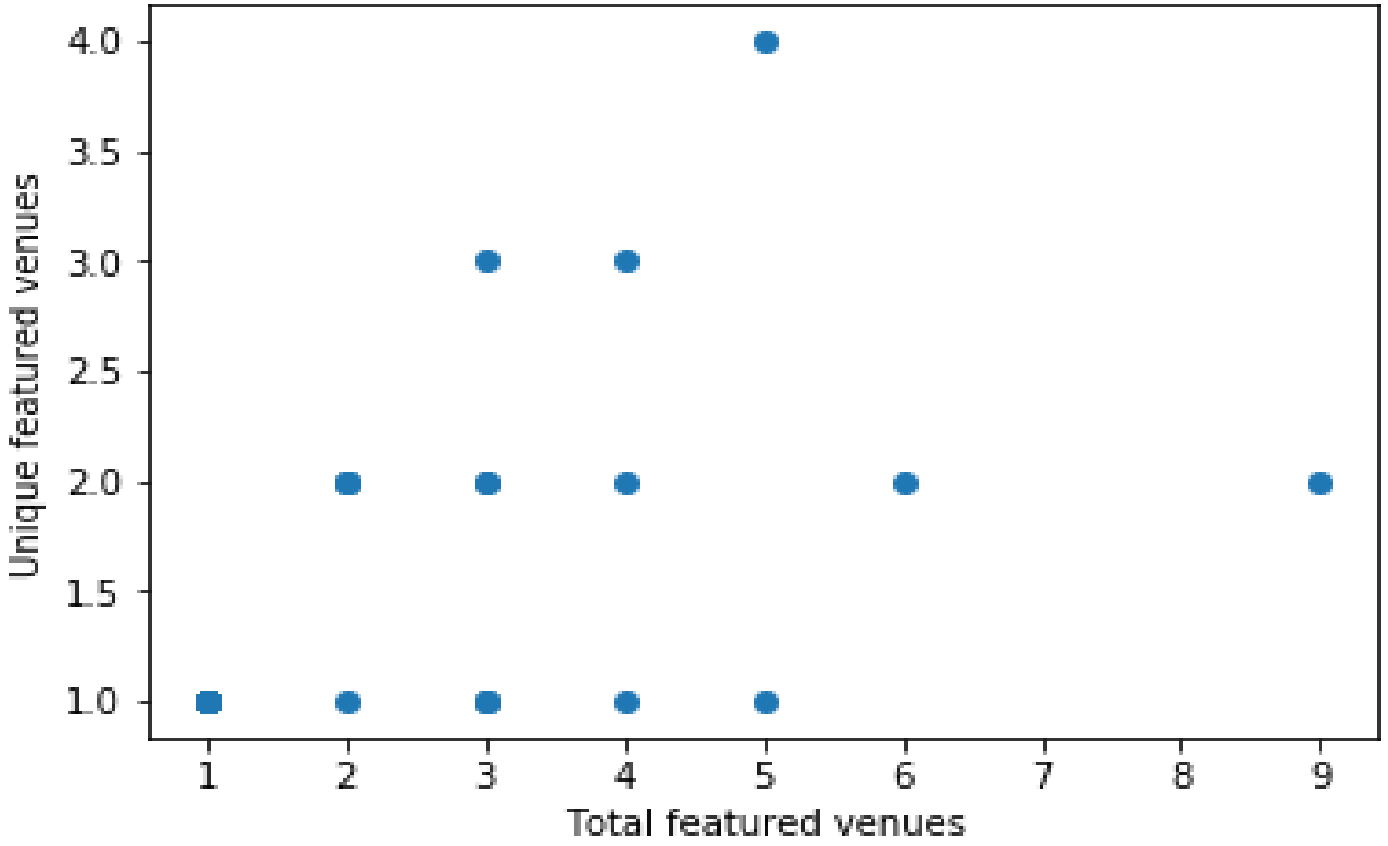
Results

- Shortlisted neighborhoods with total number of venues and total number of unique venues

	Neighborhood	Total	Unique_Total
0	Arverne	3	2
2	Astoria Heights	1	1
4	Bay Terrace	2	2
7	Beechhurst	1	1
9	Belle Harbor	3	1
11	Bilssville	1	1
12	Breezy Point	3	1
17	College Point	1	1
18	Corona	1	1
21	Edgemere	4	3
23	Far Rockaway	2	2
26	Forest Hills	2	1
27	Forest Hills Gardens	1	1
31	Hammels	5	1
33	Hollis	4	2
36	Hunters Point	5	4
38	Jamaica Center	1	1
43	Laurelton	1	1
44	Lefrak City	1	1
49	Maspeth	1	1
50	Middle Village	1	1
52	Neponsit	4	1
55	Ozone Park	1	1
56	Pomonok	3	3
58	Queensboro Hill	1	1
59	Queensbridge	2	2
65	Rockaway Beach	9	2
66	Rockaway Park	6	2
68	Roxbury	1	1
69	Somerville	1	1
71	South Ozone Park	3	1
72	Springfield Gardens	1	1
73	St. Albans	1	1
75	Sunnyside	1	1
79	Woodhaven	3	2
80	Woodside	1	1

Results

- Scatter plot of total number of unique venues vs total number of venues for each shortlisted neighborhood



- 13 points instead of 36 as there are some repeated coordinates

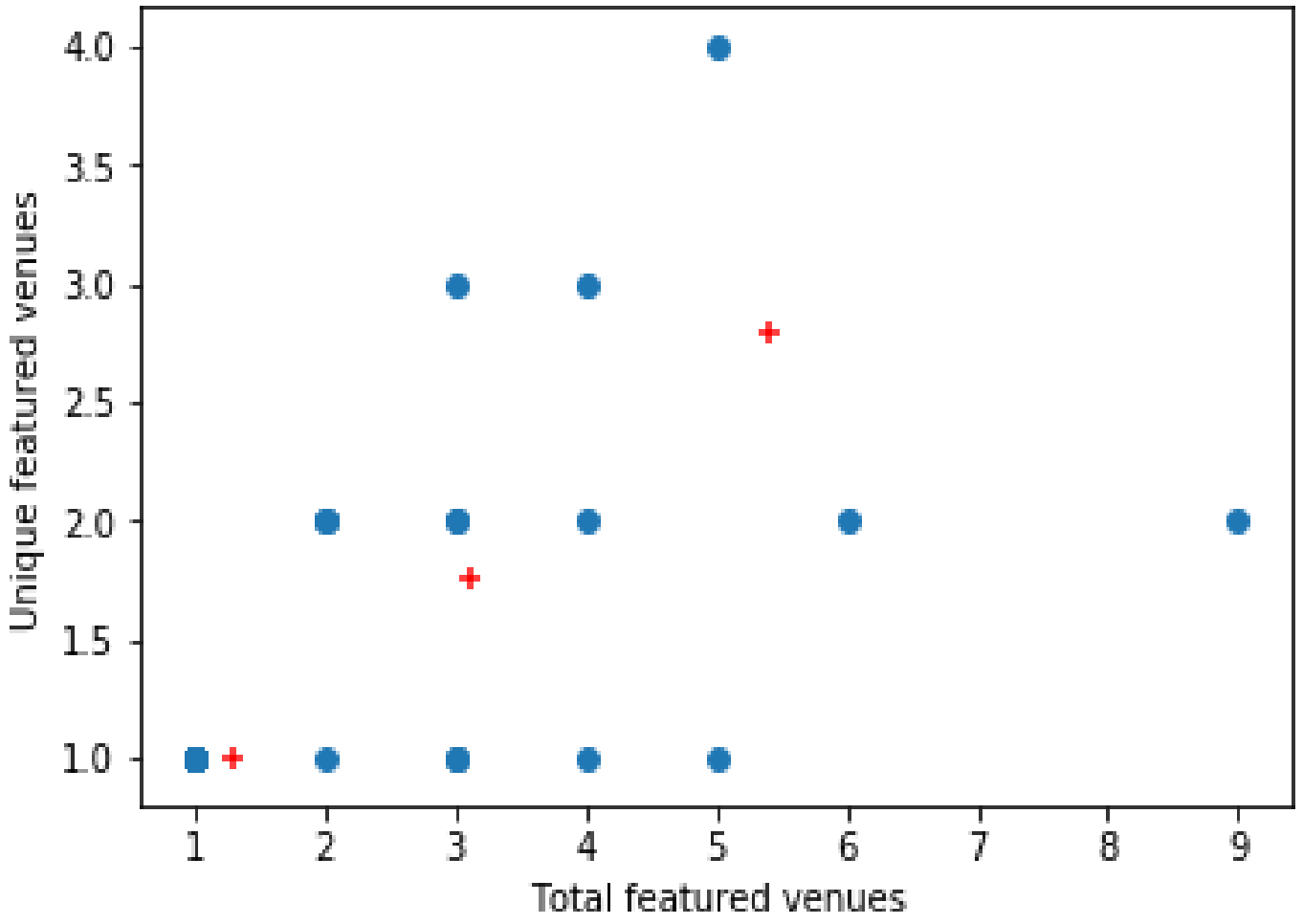
Results

- K-means clustering: To identify the ideal location for Indian fast food joint from the 36 shortlisted neighborhoods.
- Number of clusters: 4
- Clusters with their mean values

Labels	Total	Unique_Total
0	3.090909	1.545455
1	4.000000	3.333333
2	1.050000	1.000000
3	7.500000	2.000000

Results

- Scatter plot of neighborhoods and centroid of four clusters



Results

- Shortlisted neighborhoods sorted by cluster

	Neighborhood	Total	Unique_Total	Labels
0	Arverne	3	2	0
79	Woodhaven	3	2	0
33	Hollis	4	2	0
31	Hammels	5	1	0
59	Queensbridge	2	2	0
71	South Ozone Park	3	1	0
23	Far Rockaway	2	2	0
9	Belle Harbor	3	1	0
4	Bay Terrace	2	2	0
12	Breezy Point	3	1	0
52	Neponsit	4	1	0
21	Edgemere	4	3	1
36	Hunters Point	5	4	1
56	Pomonok	3	3	1
72	Springfield Gardens	1	1	2
69	Somerville	1	1	2
55	Ozone Park	1	1	2
68	Roxbury	1	1	2
73	St. Albans	1	1	2
58	Queensboro Hill	1	1	2
75	Sunnyside	1	1	2
43	Laurelton	1	1	2
45	Maspeth	1	1	2
44	Lefrak City	1	1	2
38	Jamaica Center	1	1	2
27	Forest Hills Gardens	1	1	2
26	Forest Hills	2	1	2
18	Corona	1	1	2
17	College Point	1	1	2
11	Bilssville	1	1	2
7	Beechhurst	1	1	2
2	Astoria Heights	1	1	2
50	Middle Village	1	1	2
80	Woodside	1	1	2
65	Rockaway Beach	9	2	3
66	Rockaway Park	6	2	3

Discussion

- There are total 81 neighborhoods in Queens, New York.
- The bar plot of these 81 neighborhoods vs number of venues indicated ‘Astoria’ and ‘Sunnyside Gardens’ as potential neighborhood for Indian fast food joint as number of venues in those neighborhoods are maximum.
- But as both these neighborhoods have Indian restaurant, they are removed from the list.
- 68 neighborhoods are shortlisted which included those neighborhoods which do not have Indian restaurants.
- Out of these 68 neighborhoods, only 36 neighborhoods are retained as others do not have any featured venue.
- These 36 shortlisted neighborhoods are clustered using k-means clustering approach to create 4 clusters.

Discussion

- Cluster ‘2’: Mean value of 1.05 for total featured venues, 1 for total number of unique featured venues. As this cluster has small value total featured venues as well as unique featured venues, it is not considered as ideal cluster for Indian fast food joint.

- Cluster ‘0’ has mean value of 3.09 for total featured venues and 1.54 for total number of unique featured venues. As this cluster has small value for unique featured venues, it is not considered as ideal cluster for Indian fast food joint.

- Cluster ‘3’ has mean value of 7.5 for total featured venues and 2 for total number of unique featured venues. This cluster has very high value for unique featured venues but has less value for unique featured venues. Hence, it is not considered as ideal cluster for Indian fast food joint

- **Cluster ‘1’** has mean value of 4 for total featured venues and 3.33 for total number of unique featured venues. As this cluster has high value for total featured venues as well as unique featured venues, it is an **ideal cluster for Indian fast food joint.**

Conclusion

- Determination of correct location is very important for any new restaurant. The location has major role to play for a successful restaurant.
- The aim of this project is to find ideal location for Indian fast food joint in Queens, New York.
- After retrieving the data of neighborhoods from Queens, it is filtered based on absence of Indian restaurant and presence of ‘Beach’, ‘Bowling Alley’, ‘Flea Market’, ‘Metro station’, ‘Shopping Mall’, ‘Multiplex, ‘Park’, ‘Scenic Lookout’ and ‘Movie Theater.
- The shortlisted list is then clustered using k-means clustering approach into 4 clusters. The cluster with more number of total featured venues and total unique featured venues is selected as ideal cluster for the ideal location.
- This cluster includes ‘Edgemere’, ‘Hunters Point’ and Pomonok.
- The ‘Hunters Point’ neighborhood is selected as the ideal location for Indian fast food joint based on its score on total featured venues and total unique featured venues.

Future Scope

- The approach of finding the ideal location for Indian fast food joint is based 9 features in this project.
- More number of features can be included such as ‘Hostel’, ‘College Academic Building’, ‘Vegetarian/Vegan restaurant’ etc.