

# PROSPECTS OF NEW RESTAURANT BUSINESS IN LONDON

**Coursera Capstone Project -**  
*Battle of Neighbourhoods*

**Vidya Rani Puliyangodan Kallaiveedu**



# TABLE OF CONTENTS

- 1. Introduction**
- 2. Data**
- 3. Methodology - Exploratory Data Analysis**
- 4. Results**
- 5. Discussion**
- 6. Conclusion**

# INTRODUCTION

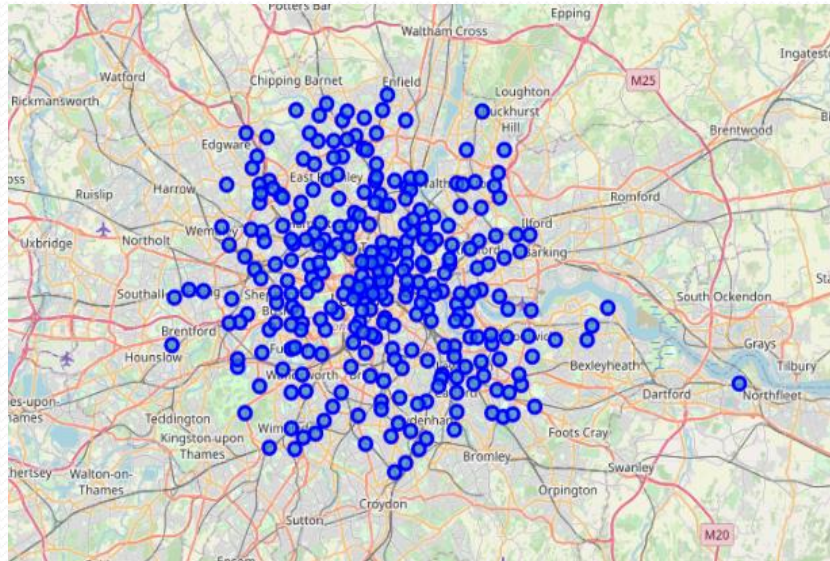
- **The business problem we are trying to solve** is – If an investor is interested in opening a new eating spot in the capital city London –the model should be capable of providing insight on
  - what would be the best profitable type of restaurant for a given location?
  - Or what would be the best location for a specific type of cuisine?
- **Target audience** is - entrepreneurs or investors considering to start a new restaurant or franchise requiring guidance in picking the locality for a particular gastronomy OR cuisine for definite location

# DATA

- London boroughs and locations data from [https://en.wikipedia.org/wiki/List\\_of\\_areas\\_of\\_London](https://en.wikipedia.org/wiki/List_of_areas_of_London) (Wikipedia). Locations with “Post Town” – London are selected for analysis.
- Use the Geopy and Folium library to get the coordinates of every location and map geospatial data on a London map.
- Foursquare API is used to collect the top 200 restaurants and their categories for each location within a radius 500 meters.

# Methodology - Exploratory Data Analysis

- **EXPLORING LONDON NEIGHBORHOODS**
  - London borough and Neighborhood information is scraped from Wikipedia
  - Using Geospatial data – Longitude and latitude of the neighborhood are populated
  - Final dataframe has - **32 boroughs and 299 neighbourhoods**



## • EXPLORING LONDON RESTAURANTS

- Foursquare API is used to explore the neighborhoods and segment them by fetching top 200 restaurants in each neighbourhood
- The results have 3342 unique restaurants in London with 88 different styles of cuisines.
- one-hot encoding is applied to group them by taking the mean of the frequency of occurrence of each type.
- Clustered neighborhoods using the k-means algorithm based on their types similarity.
- Best K is calculated for the data . i.e. K=0

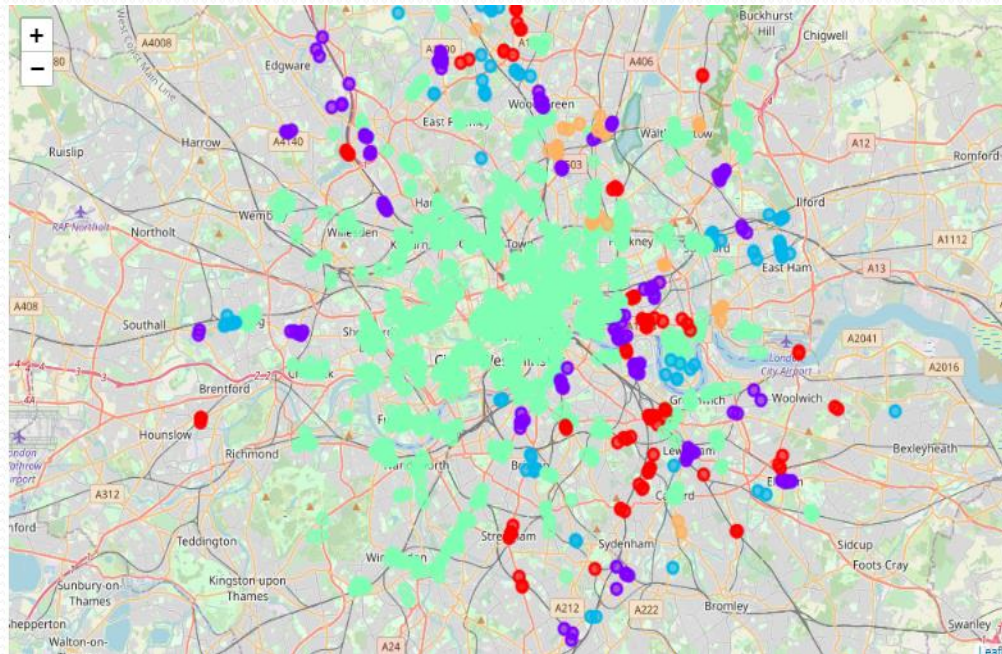
	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
2	Acton	51.50814	-0.273261	Amigo's Peri Peri	51.508396	-0.274561	Fast Food Restaurant	1	Fast Food Restaurant	Chinese Restaurant	Restaurant
4	Acton	51.50814	-0.273261	North China Restaurant	51.508251	-0.277435	Chinese Restaurant	1	Fast Food Restaurant	Chinese Restaurant	Restaurant
5	Acton	51.50814	-0.273261	Subway	51.507501	-0.271765	Restaurant	1	Fast Food Restaurant	Chinese Restaurant	Restaurant
6	Acton	51.50814	-0.273261	Ming's	51.507456	-0.272260	Chinese Restaurant	1	Fast Food Restaurant	Chinese Restaurant	Restaurant
9	Acton	51.50814	-0.273261	Sam's Chicken	51.507193	-0.270431	Fast Food Restaurant	1	Fast Food Restaurant	Chinese Restaurant	Restaurant



# RESULTS

- All 5 clusters follow unique pattern for top ten common restaurants for a particular neighborhood.

Clustered map is shown below



# Cluster 1

	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
count	120.0	120	120	120	120	120	120	120	120	120	120
unique	NaN	10	10	13	12	11	10	9	9	8	8
top	NaN	Chinese Restaurant	Chinese Restaurant	Indian Restaurant	Yoshoku Restaurant	Halal Restaurant	Falafel Restaurant	Fast Food Restaurant	Falafel Restaurant	Fast Food Restaurant	Filipino Restaurant
freq	NaN	60	32	22	21	18	30	29	27	27	27

- The most common restaurant: Chinese Restaurant
- 10th most common restaurant: Filipino Restaurant



## Cluster 2

	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
count	177.0	177	177	177	177	177	177	177	177	177	177
unique	NaN	2	19	14	15	13	12	12	7	7	7
top	NaN	Fast Food Restaurant	Italian Restaurant	Restaurant	Italian Restaurant	Chinese Restaurant	Portuguese Restaurant	Filipino Restaurant	Gluten-free Restaurant	Filipino Restaurant	French Restaurant
freq	NaN	165	22	42	27	25	37	31	38	40	40

- The most common restaurant: Fast Food Restaurant
- 10th most common restaurant: French Restaurant

## Cluster 3

	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
count	86.0	86	86	86	86	86	86	86	86	86	86
unique	NaN	6	9	8	11	7	7	6	5	4	4
top	NaN	Indian Restaurant	Fast Food Restaurant	Yoshoku Restaurant	Halal Restaurant	Ethiopian Restaurant	Falafel Restaurant	Filipino Restaurant	French Restaurant	German Restaurant	Gluten-free Restaurant
freq	NaN	72	22	28	26	28	37	31	31	31	31

- The most common restaurant: Indian Restaurant
- 10th most common restaurant: Gluten-free Restaurant

## Cluster 4

	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
count	2899.0	2899	2899	2899	2899	2899	2899	2899	2899	2899	2899
unique	NaN	28	31	39	40	39	44	50	44	47	52
top	NaN	Italian Restaurant	Italian Restaurant	Restaurant	Indian Restaurant	Restaurant	Sushi Restaurant	Thai Restaurant	Seafood Restaurant	Falafel Restaurant	Mexican Restaurant
freq	NaN	1188	727	352	404	308	227	203	262	168	220

- The most common restaurant: Italian Restaurant
- 10th most common restaurant: Mexican Restaurant

## Cluster 5

	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
count	60.0	60	60	60	60	60	60	60	60	60	60
unique	NaN	2	6	6	6	7	7	6	4	5	5
top	NaN	Turkish Restaurant	Mediterranean Restaurant	Kebab Restaurant	Restaurant	Mexican Restaurant	Middle Eastern Restaurant	Portuguese Restaurant	Falafel Restaurant	Cuban Restaurant	Caribbean Restaurant
freq	NaN	53	22	22	22	22	22	22	32	22	22

- The most common restaurant: Turkish Restaurant
- 10th most common restaurant: Caribbean Restaurant

# Summary of results

Cluster	Most Common Restaurant	Least Common Restaurant
1	Chinese Restaurant	Filipino Restaurant
2	Fast Food Restaurant	French Restaurant
3	Indian Restaurant	Gluten-free Restaurant
4	Italian Restaurant	Mexican Restaurant
5	Turkish Restaurant	Caribbean Restaurant

# DISCUSSION

- Safest option to choose the restaurant type for a particular locality is by considering the least common type.

Opening an Indian restaurant in a street filled with dozens of Indian dining is a risky investment.

- Choosing the least popular restaurant is also unreliable, given the demand and popularity for that type of food is less in that area.

Analysis is limited to 10 common restaurants to eliminate the risk of opening a business with no prospect in a location



## Recommended Cuisines for each cluster

Cluster	Least Common Restaurants
1	Filipino Restaurant or Fast Food Restaurant
2	French Restaurant or Filipino Restaurant
3	Gluten-free Restaurant or German Restaurant
4	Mexican Restaurant or Falafel Restaurant
5	Caribbean Restaurant or Cuban Restaurant

# CONCLUSION

- The analysis is performed to find out the most promising category of restaurant and appropriate location to start the business.
- The data is collected from Wikipedia, geospatial library and foursquare API. The method used is K-means clustering to group the similar neighbourhood on the basis of frequency of certain cuisine types.
- Undoubtedly the model can be improved by collecting further information about the neighbourhood population (/demography), restaurant menu and ratings to come up with improved recommendation.
- The analysis can be applied to any city of your choice that has available geospatial information.
- The analysis can be extended to any kind of venue (shopping, clubs, etc.) that is available Foursquare database.