

# The Battle of Neighbourhoods Final

Presented by : Vishal Chavan

# 1. Introduction

---

**Background:** All the people over the world love to dine in different restaurants. Here, we would like to set up new Indian Restaurant in London with multiple varieties.

**Problem:** This Project aim to select the borough in London based on their popularity, explore the neighborhoods of that borough to find the common venues in each neighborhood and finally cluster the neighborhoods using K-means clustering.

**Target:** Considering the diversity in London where most people are multireligion based. London is a place where different shades live. Definitely, by looking at the population we can determine there is highly shortage of Indian restaurants in London.

## 2. Data acquisition and cleaning:

---

Data Acquisition: The Data acquired for this project is a combination of data from 2 different sources:

- The first data source of the project used a London area that shows data of borough in London.
- The second source of data is demographic of London where we can see demography of the people living in London.

Data Cleaning: The Data cleaning process for each source of data are done separately.

- From London list of areas, it will show London borough with postal town and postal code which are then pivoted to check total numbers.
- Demography of London where we have sorted white, black, mixed, Asian to check how much count in each location in London
- Latitude and Longitude are obtained and join with location which will help us to identify the distance

# 3. Methodology

---

Exploring Data Analysis:

The Describe function in python is used to get statistic of the London areas, this return the mean, standard deviation, minimum, maximum, 1<sup>st</sup> Quantile (25%), 2<sup>nd</sup> Quantile (50%), 3<sup>rd</sup> Quantile (75%) for each of the major categories

Count in Diagram shows venue unique counts describe in categories of venues.

	Count
<b>count</b>	195.000000
<b>mean</b>	17.600000
<b>std</b>	38.354881
<b>min</b>	1.000000
<b>25%</b>	3.000000
<b>50%</b>	6.000000
<b>75%</b>	17.000000
<b>max</b>	347.000000

# Methodology Cont.....

---

## Modelling:

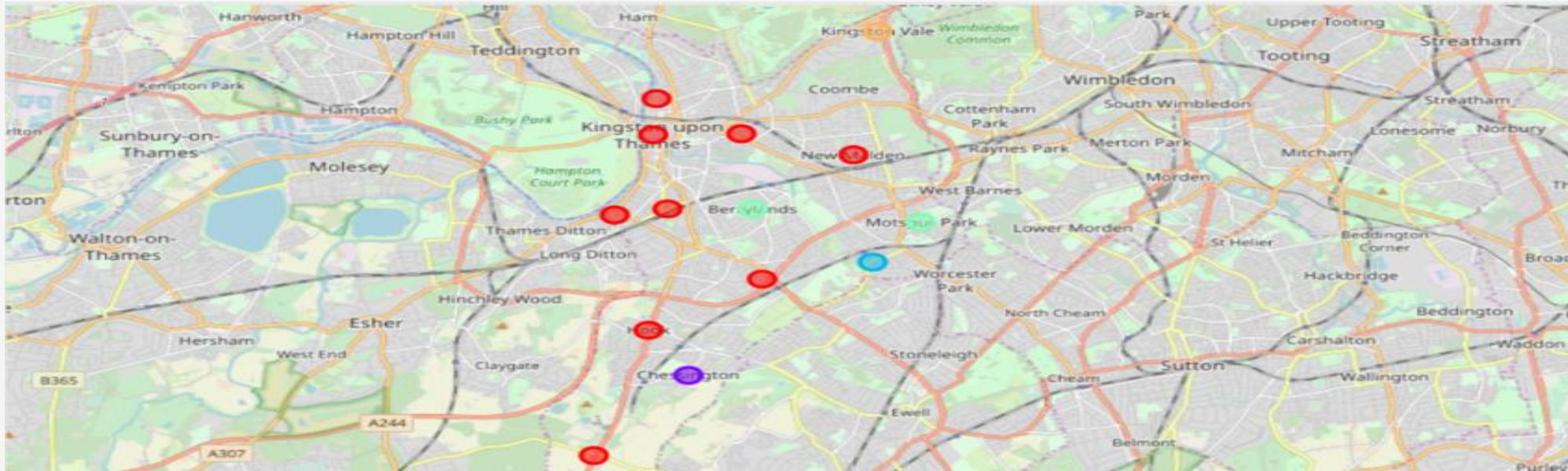
- Using Final data set containing the neighbourhood in London with latitude and longitude, we can find all venues within a 500meter radius of each neighbourhood by connecting to the Foursquare API
- One hot encoding is done on the venues data. The venues data is then grouped by neighbourhood and the mean of the venues are calculated, finally 10 common venues are calculated for each neighbourhoods.
- To help people find similar neighbourhoods in the borough we will be clustering similar neighbourhoods using K –means clustering which is a form of unsupervised machine learning algorithm that clusters data based on predefined cluster size

	Neighbourhood	Neighbourhood Latitude	Neighbourhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Crofton Park	51.46268	-0.03558	The Orchard	51.463678	-0.035699	Gastropub
1	Crofton Park	51.46268	-0.03558	Brockley's Rock	51.459457	-0.033868	Fish & Chips Shop
2	Crofton Park	51.46268	-0.03558	Browns Of Brockley	51.464513	-0.037346	Coffee Shop
3	Crofton Park	51.46268	-0.03558	Waterintobeer	51.463712	-0.038826	Beer Store
4	Crofton Park	51.46268	-0.03558	Saka Maka	51.464826	-0.036437	Indian Restaurant

## 4.Results

---

After running the K-means clustering we can access each cluster created to see which neighbourhoods were assigned to each of the five cluster.



To find the optimal value of the number of clusters,  $k$ , the number of clusters is iterated corresponding Silhouette Coefficient is calculated for each of the  $k$ -values used.

# Results Cont.....

Cluster 1:

	Borough	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
9	Lewisham	0	Pub	Grocery Store	Park	Café	Supermarket	Gym / Fitness Center	Italian Restaurant	Indian Restaurant	Train Station
15	Lewisham	0	Pub	Grocery Store	Park	Café	Supermarket	Gym / Fitness Center	Italian Restaurant	Indian Restaurant	Train Station
20	Lewisham	0	Grocery Store	Park	Supermarket	Café	Pub	Italian Restaurant	Coffee Shop	Fast Food Restaurant	Train Station
26	Lewisham	0	Grocery Store	Park	Supermarket	Café	Pub	Italian Restaurant	Coffee Shop	Fast Food Restaurant	Train Station
36	Lewisham	0	Grocery Store	Park	Supermarket	Café	Pub	Italian Restaurant	Coffee Shop	Fast Food Restaurant	Train Station
37	Lewisham	0	Pub	Grocery Store	Park	Café	Supermarket	Gym / Fitness Center	Italian Restaurant	Indian Restaurant	Train Station

Upon looking these neighborhoods we can see most common venues in these neighborhoods are pubs and Restaurants

# Results Cont.....

Cluster 2:

	Borough	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
0	Lewisham	1	Coffee Shop	Pub	Café	Park	Gastropub	Bar	Indian Restaurant	Italian Restaurant	Cocktail Bar
2	Lewisham	1	Pub	Café	Coffee Shop	Bar	Park	Brewery	Trail	Indie Movie Theater	Italian Restaurant
4	Southwark	1	Pub	Café	Pizza Place	Italian Restaurant	Park	Coffee Shop	Gastropub	Burger Joint	Garden Center
11	Lewisham	1	Pub	Café	Gastropub	Park	Garden	Food Truck	Fish & Chips Shop	Coffee Shop	Restaurant
13	Lewisham	1	Coffee Shop	Pub	Café	Park	Gastropub	Bar	Indian Restaurant	Italian Restaurant	Cocktail Bar
16	Lewisham	1	Pub	Café	Gastropub	Park	Garden	Food Truck	Fish & Chips Shop	Coffee Shop	Restaurant
17	Lewisham	1	Pub	Coffee Shop	Café	Bar	Italian Restaurant	Park	Indian Restaurant	Gastropub	Brewery
19	Southwark	1	Pub	Café	Coffee Shop	Pizza Place	Park	Bar	Gastropub	Burger Joint	Cocktail Bar
21	Southwark	1	Pub	Café	Coffee Shop	Pizza Place	Park	Bar	Gastropub	Burger Joint	Cocktail Bar
22	Southwark	1	Pub	Brewery	Coffee Shop	Park	Bar	Café	Food Truck	Gym / Fitness Center	Vietnamese Restaurant

The second cluster has neighbourhoods which consists of venue CAFÉ, Pubs, Park



# Results Cont.....

Cluster 3:

	Borough	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
1	Southwark	2	Café	Coffee Shop	Park	Pub	Italian Restaurant	Cocktail Bar	Middle Eastern Restaurant	Brewery	Pizza Place
10	Lambeth	2	Coffee Shop	Pub	Café	Brewery	Market	Pizza Place	Cocktail Bar	Burger Joint	Caribbean Restaurant
29	Lambeth	2	Coffee Shop	Pub	Café	Brewery	Market	Pizza Place	Cocktail Bar	Burger Joint	Caribbean Restaurant
35	Southwark	2	Café	Coffee Shop	Park	Pub	Italian Restaurant	Cocktail Bar	Middle Eastern Restaurant	Brewery	Pizza Place

The third cluster has neighbourhoods which consists of venue Coffee shops, Pubs.

# Results Cont.....

Cluster 4:

	Borough	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
3	Southwark	3	Pub	Café	Park	Coffee Shop	Grocery Store	Bakery	Gym / Fitness Center	Farmers Market	Pizza Place
7	Lewisham	3	Pub	Coffee Shop	Café	Park	Grocery Store	Supermarket	Gym / Fitness Center	Japanese Restaurant	Pizza Place
3	Lambeth	3	Pub	Park	Coffee Shop	Italian Restaurant	Grocery Store	Café	Train Station	Gastropub	Bakery
12	Lewisham	3	Pub	Coffee Shop	Café	Park	Grocery Store	Supermarket	Gym / Fitness Center	Japanese Restaurant	Pizza Place
23	Croydon	3	Pub	Grocery Store	Café	Supermarket	Coffee Shop	Park	Tram Station	African Restaurant	Plaza
25	Croydon	3	Pub	Grocery Store	Café	Supermarket	Coffee Shop	Park	Tram Station	African Restaurant	Plaza
30	Croydon	3	Pub	Park	Coffee Shop	Italian Restaurant	Grocery Store	Café	Train Station	Gastropub	Bakery
32	Lewisham	3	Pub	Grocery Store	Park	Coffee Shop	Garden	Bakery	Italian Restaurant	Supermarket	Clothing Store
33	Lambeth	3	Pub	Grocery Store	Coffee Shop	Café	Bakery	Park	Pizza Place	Gym / Fitness Center	Italian Restaurant

The fourth cluster has neighbourhoods which consists of venue Pubs, Park, Grocery Stores.

# Results Cont.....

Cluster 5:

	Borough	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
5	Southwark	4	Coffee Shop	Hotel	Pub	Italian Restaurant	Theater	Seafood Restaurant	Gym / Fitness Center	Art Museum	Garden
6	Southwark	4	Coffee Shop	Hotel	Pub	Italian Restaurant	Theater	Seafood Restaurant	Gym / Fitness Center	Art Museum	Garden
14	Lambeth	4	Coffee Shop	Hotel	Pub	Italian Restaurant	Theater	Seafood Restaurant	Gym / Fitness Center	Art Museum	Garden
18	Southwark	4	Coffee Shop	Hotel	Pub	Italian Restaurant	Theater	Seafood Restaurant	Gym / Fitness Center	Art Museum	Garden
24	Southwark	4	Coffee Shop	Hotel	Pub	Italian Restaurant	Theater	Seafood Restaurant	Gym / Fitness Center	Art Museum	Garden

- The fifth cluster has neighbourhoods which consists of venue Coffee Shop, Pub, Hotel

## 5. Discussion and Conclusion:

---

- It's been visible from cluster 3 and 4 that Indian Restaurant can play a vital role in Restaurant competition and very good chance to set up.
- Their proximity to other amenities and accessibility to station are huge. These 2 clusters do not have top restaurants that could rival their standards if they are created.
- The chances to resources needed is quite high as Lewisham and Lambeth. In conclusion, this project would have had better results if there were more data in terms of per capita income data within the area, traffic access, corporates of more venues exploration with the Foursquare.
- Getting the ratings and feedbacks of the current restaurants within the clusters would have helped in providing more insight into the best location

---

# Thank You