

# **IBM Data Science Capstone Project**

## **The Battle of the Neighborhoods**

### **1. Business Problem**

An Afghan business man would like to open an afghan restaurant in Greater London. There are a lot of Afghan immigrants in London that like Afghan Food. Also some people from Iran, Pakistan and Arab citizens who live in London are interested in Afghan Food. It seems that Greater London is great place for launching an afghan restaurant potentially.

Greater London has:

- 33 districts
- 32 boroughs
- the City of London

Therefore, it is highly crucial to know what place is the best to open an Afghan restaurant.

#### **Target Audience**

Customer is a business man that want to establish an afghan restaurant. I help him to make the best choice, and recommend the right borough and explain why that place is better than others.

#### **Success Criteria**

The success criteria of the project will be a good recommendation of borough for a restaurant.

### **2. Data section**

The following data sources will be used in this project:

I. Foursquare API.

a. <https://foursquare.com/>

This service is used to collect Foursquare location data, explore each borough of London and acquire information about venues, especially restaurants, café, bars, pubs, etc. I will use this data

to understand how many restaurants are in each borough. Also it will help to create map.

## II. The London Data store

- a. <https://data.london.gov.uk/>
- b. <https://data.london.gov.uk/dataset/land-area-and-population-density-ward-and-borough>

The London Data store has been created by the Greater London Authority. It has a lot of free data for every year about boroughs, populations and squares in Greater London. The most useful for me is a list of borough names and borough population for 2020 year.

## III. TripAdvisor

- a. <https://www.tripadvisor.com/>

Unfortunately, foursquare does not have information about all Afghan restaurants and Afghan Food, therefore I need to collect and check some information manually, particularly on TripAdvisor.

## 3. Methodology section

### **Business Introduction:**

The main goal is to find optimum borough for a new restaurant.

### **Analytic approach:**

Bellow two approaches have been used:

#### 1<sup>st</sup> Approach:

my goal is to find boroughs that have more restaurants and is center of attraction. At the same time, I want to exclude places where there are Afghani restaurants. I don't want to compete with Afghani restaurants. But, I believe we can attract a part of customers where there are a lot of restaurants with other kind of dishes:

- Find Top 5 boroughs that have maximum of restaurants, café, pubs in it
- Find boroughs with Afghani restaurants;
- Delete from Top 5 boroughs with Afghani restaurants.

## 2<sup>nd</sup> Approach:

My goal is to select clusters that are similar with boroughs with Afghani restaurants. If in one of boroughs, there is Afghani restaurant, that continues to work and has a lot of good reviews, then I find borough with similar infrastructure.

So, I will use k-means algorithm to cluster boroughs.

### Exploratory Data Analysis:

1. Download csv-file from <https://data.london.gov.uk/dataset/land-area-and-population-density-ward-and-borough>
2. Choose data for 2020 year.

	Code	Borough	Population	Square_km
0	E09000001	City of London	8046	2.9
1	E09000002	Barking and Dagenham	218367	36.1
2	E09000003	Barnet	407148	86.7
3	E09000004	Bexley	254970	60.6
4	E09000005	Brent	343853	43.2
5	E09000006	Bromley	336395	150.1
6	E09000007	Camden	257384	21.8
7	E09000008	Croydon	400272	86.5
8	E09000009	Ealing	361833	55.5
9	E09000010	Enfield	342993	80.8
10	E09000011	Greenwich	292466	47.3
11	E09000012	Hackney	289488	19.0
12	E09000013	Hammersmith and Fulham	190721	16.4

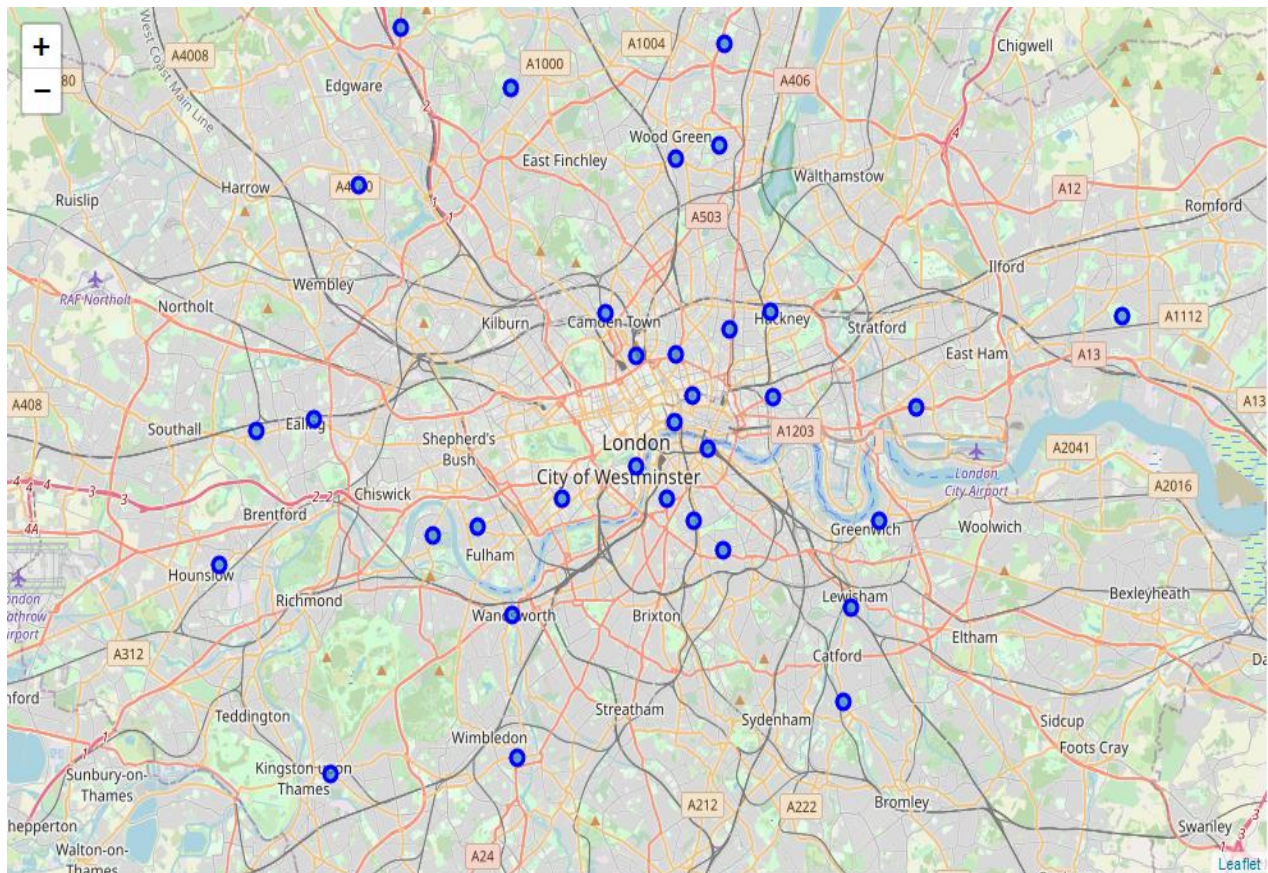
*Fig. 1. borough list from The London Data store.*

Use python geocoder library, find Latitude and Longitude, and add data to list of boroughs.

	Code	Borough	Population	Square_km	Latitude	Longitude
0	E09000001	City of London	8046	2.9	51.520500	-0.097430
1	E09000002	Barking and Dagenham	218367	36.1	51.543932	0.133157
2	E09000003	Barnet	407148	86.7	51.627294	-0.253759
3	E09000004	Bexley	254970	60.6	51.622832	-0.080656
4	E09000005	Brent	343853	43.2	51.609776	-0.194679
5	E09000006	Bromley	336395	150.1	51.431820	-0.016566
6	E09000007	Camden	257384	21.8	51.532360	-0.127960
7	E09000008	Croydon	400272	86.5	51.593209	-0.083390
8	E09000009	Ealing	361833	55.5	51.514060	-0.300730
9	E09000010	Enfield	342993	80.8	51.540021	-0.077501
10	E09000011	Greenwich	292466	47.3	51.484540	0.002750
11	E09000012	Hackney	289488	19.0	51.545050	-0.055320
12	E09000013	Hammersmith and Fulham	190721	16.4	51.482690	-0.212910

*Fig. 2. A borough table with Latitude and Longitude.*

Create a map of all boroughs by folium.



*Fig. 3. A map of boroughs of Greater London*

We can see from Fig. 3 that some boroughs are far from center. So, those are not best place.

Find nearby venues for every borough by Foursquare API (radius 1000 m) and create a sorted list of boroughs with places similar to restaurants, cafes.

**Boroughs with restaurants and cafés:**

Borough	
Enfield	48
Hammersmith and Fulham	47
Lambeth	46
Ealing	43
Islington	43
Merton	43
Tower Hamlets	42
Southwark	40
Greenwich	38
Kingston upon Thames	38
Harrow	38
Havering	38
City of London	38
Redbridge	36
Camden	33
Hillingdon	32
Sutton	31
Westminster	26
Haringey	25
Wandsworth	24
Hackney	23
Brent	20
Croydon	18
Hounslow	15
Newham	15
Kensington and Chelsea	14
Richmond upon Thames	13
Lewisham	12
Waltham Forest	7
Barnet	3
Bromley	2
Bexley	2

Find Afghani restaurants. Unfortunately, Foursquare can't find all information about Afghani restaurants. Therefore, I collect information from TripAdvisor:

	<b>Name</b>	<b>Borough</b>	<b>Rating</b>
<b>0</b>	Kabul_Food	Kensington and Chelsea	8
<b>1</b>	Kolba	Kensington and Chelsea	8
<b>2</b>	Palaw Resturant	Kensington and Chelsea	10
<b>3</b>	Zam Zam	Westminster	8
<b>4</b>	Balkh_Resturant	Islington	8

*Fig. 4. A table of Afghani restaurants in Greater London*



We can see:

- Only 5 well-known Afghani restaurants in Greater London
- 3 of them are located in Kensington, one in Westminster and 1 in Islington.
- All restaurants have good rating (equals or more than 8) and reviews.

One of the goals as was to find boroughs where there are a lot of restaurants except Afghani restaurants.

Use data from 5 and 6 steps and create a short list of good boroughs for us:

Borough	
Enfield	48
Hammersmith and Fulham	47
Lambeth	46
Ealing	43
Merton	43

Find each borough from short list along with the top 5 most common venues:

----Enfield---- venue freq 0 Café 0.10 1 Pub 0.10 2 Coffee Shop 0.06 3 Restaurant 0.05 4 Cocktail Bar 0.03	----Hammersmith and Fulham---- venue freq 0 Café 0.11 1 Italian Restaurant 0.09 2 Coffee Shop 0.07 3 Pub 0.07 4 Grocery Store 0.06
----Lambeth---- venue freq 0 Pub 0.16 1 Café 0.08 2 Hotel 0.07 3 Coffee Shop 0.06 4 Park 0.05	----Ealing---- venue freq 0 Coffee Shop 0.11 1 Pub 0.11 2 Hotel 0.05 3 Park 0.04 4 Italian Restaurant 0.04
----Merton---- venue freq 0 Grocery Store 0.06 1 Coffee Shop 0.06 2 Italian Restaurant 0.05 3 Pub 0.05 4 Thai Restaurant 0.04	

We can see again these boroughs have good infrastructure if we want to use them as a center of attraction.

Create a data frame and display the top 10 venues for each borough:

	Borough	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	Barking and Dagenham	Metro Station	Convenience Store	Park	Plaza	Gift Shop	Grocery Store	Fast Food Restaurant	Film Studio	Fish & Chips Shop	Fish Market
1	Barnet	Italian Restaurant	Hotel	Pub	Sandwich Place	Coffee Shop	Fast Food Restaurant	Supermarket	Golf Course	Bookstore	Rest Area
2	Bexley	Intersection	Sandwich Place	Park	Performing Arts Venue	Miscellaneous Shop	Grocery Store	Mexican Restaurant	Coffee Shop	Pizza Place	Convenience Store
3	Brent	Café	Indian Restaurant	Japanese Restaurant	Grocery Store	Coffee Shop	Turkish Restaurant	Golf Course	Metro Station	Middle Eastern Restaurant	Supermarket
4	Bromley	Park	Grocery Store	Train Station	Hardware Store	Shipping Store	Event Service	Trail	Fast Food Restaurant	Gym	Supermarket
5	Camden	Café	Coffee Shop	Burger Joint	Park	Italian Restaurant	Breakfast Spot	Pizza Place	Art Gallery	Hotel	Mexican Restaurant
6	City of London	Coffee Shop	Gym / Fitness Center	Hotel	Vietnamese Restaurant	Pub	Food Truck	Bar	Café	Wine Bar	Whisky Bar
7	Croydon	Grocery Store	Café	Pub	Park	Bus Stop	Fast Food Restaurant	Turkish Restaurant	Supermarket	Convenience Store	Middle Eastern Restaurant
8	Ealing	Coffee Shop	Pub	Hotel	Park	Italian Restaurant	Thai Restaurant	Pizza Place	Platform	Burger Joint	Sushi Restaurant
9	Enfield	Pub	Café	Coffee Shop	Restaurant	Bakery	Cocktail Bar	Park	Middle Eastern Restaurant	Grocery Store	Lounge
10	Greenwich	Pub	Grocery Store	Café	Garden	Pizza Place	Coffee Shop	Italian Restaurant	Park	Historic Site	Trail
11	Hackney	Pub	Coffee Shop	Café	Bakery	Cocktail Bar	Restaurant	Garden	Beer Store	Butcher	Wine Shop

Fig. 5. A fragment of table the top 10 venues for each borough.

Run k-means clustering, find 4 similar clusters of boroughs and create a map.

Light blue circles are the best places for Afghani restaurant. Boroughs such as Kensington, Chelsea, Westminster and Islington have been excluded. because we don't want to compete with existent Afghan restaurants.

If we compare a list from step 5 and Fig. 6 we can see Lambeth and Tower Hamlets are the best boroughs for Afghani restaurant.



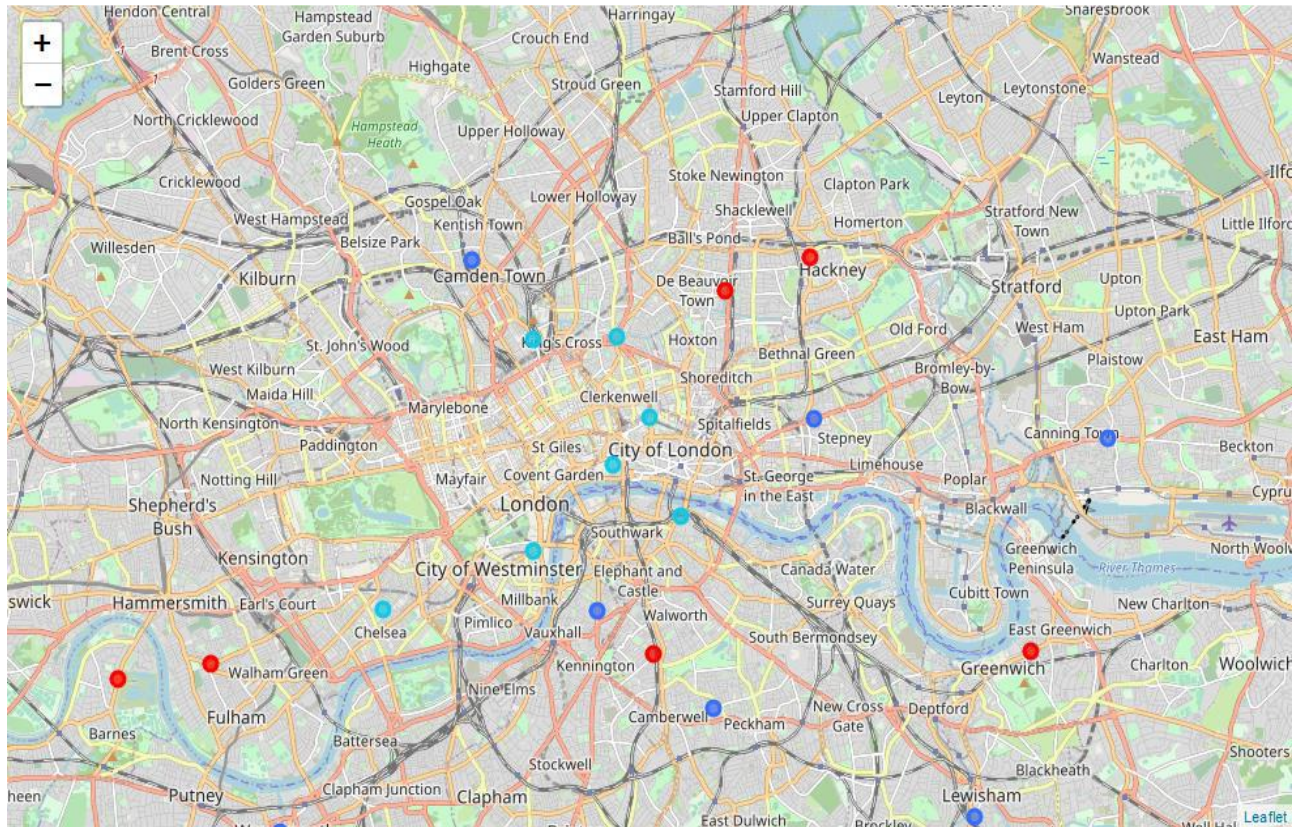


Fig. 6. A map of similar borough clusters in Greater London

#### 4. Results

Main research results:

- I. The best boroughs to open an Afghan restaurant are Lambeth and Tower Hamlets;
- II. Greater London don't have a lot of Afghan restaurant;
- III. Existent Afghani restaurants have relatively good location: Kensington and Chelsea, Westminster and Islington and they have high rating;
- IV. Lambeth and Tower Hamlets boroughs are similar to boroughs where current Afghani restaurants are located.

## **5. Discussion**

Lambeth and Tower Hamlets are the best boroughs to open Afghani restaurant.

These boroughs have more restaurants, cafes and entertainments centers that can be used as 'a center of attraction'. Likewise, these boroughs don't have direct competitors (Afghani restaurant). It's important for business development.

I avoid boroughs such as Kensington, Chelsea, Westminster and Islington to run an Afghani restaurant. because they have Afghani restaurants that all of are of high ratings and reviews.

## **6. Conclusion**

The aim was to find the best borough in Greater London to run an Afghani restaurant.

Bellow criteria's have been taken into consideration:

- a. Center of attraction
- b. Minimum Afghan restaurants
- c. k-means clustering to find similar boroughs where Afghan restaurants are located now.

As a result, we can recommend are Lambeth and Tower Hamlets boroughs.

In the future, it's a good idea to collect more information about revenue, population, rent payments and crime level for future projects.