

Finding the optimal location for an Indian restaurant in London

By Ismael Siffon

Problem description

In this project, we are going to help a restaurateur pick the location of his new Indian restaurant in London. With the increasing rate at which Indian cuisine is becoming popular in the UK, our restaurateur has some specific requirements for his location:

- **Safety:** high crime should be avoided.
- **Costumer profile:** high Indian population density, or other nationalities that have similar cuisine.
- **Competition:** should not be crowded with Indian restaurants.
- **Customer purchasing power:** middle-class income.
- **Location affordability:** Rent should be affordable.

Data gathering and cleaning

Demographic data of all boroughs in London:

- Population density
- Average age
- Largest, second and third migrant population and percentage
- Employment rate
- Gross annual pay
- Crime rate, among others.

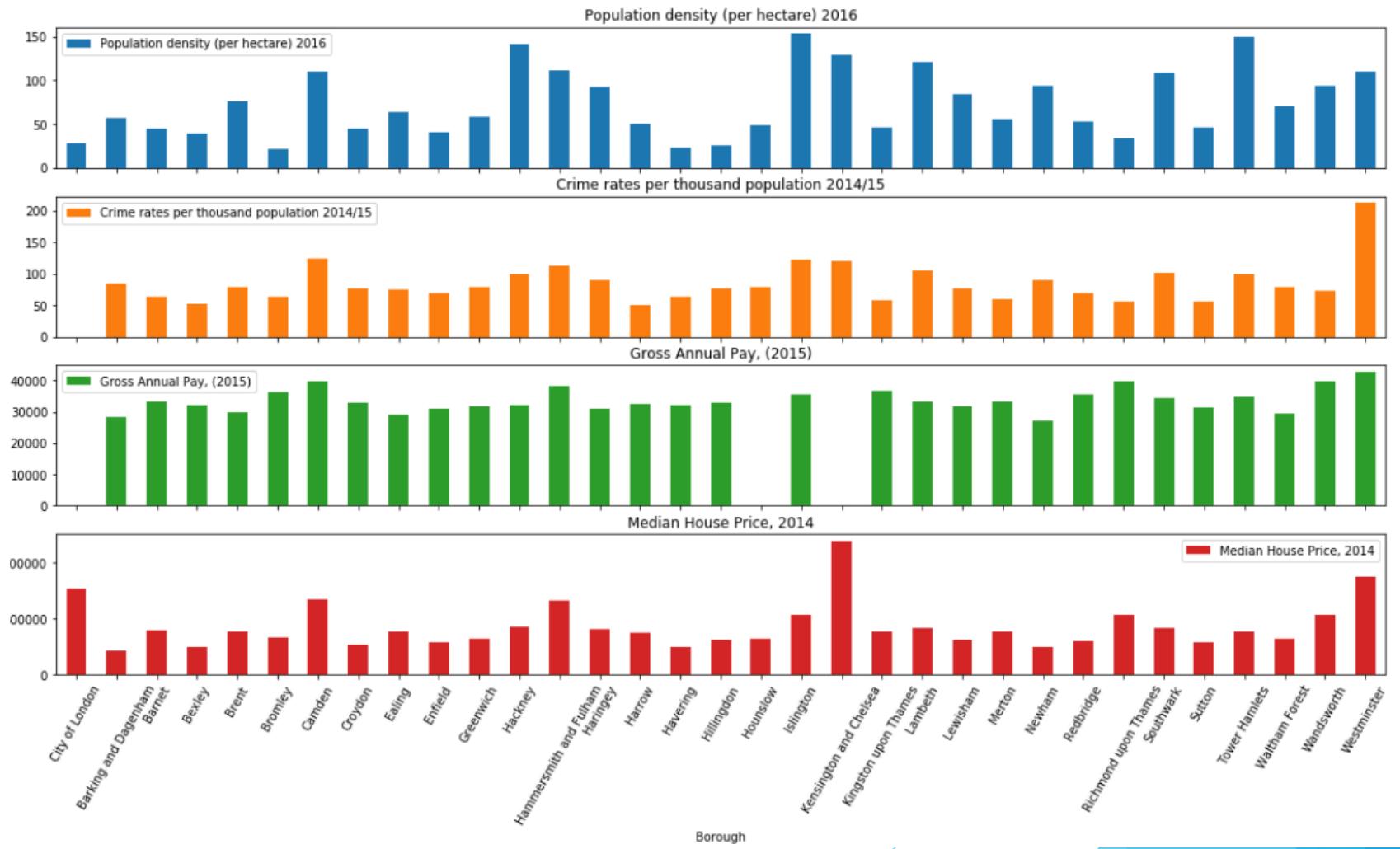
List of neighborhoods and their coordinates:

- Web-scraping from Wikipedia
- Google API

Foursquare data:

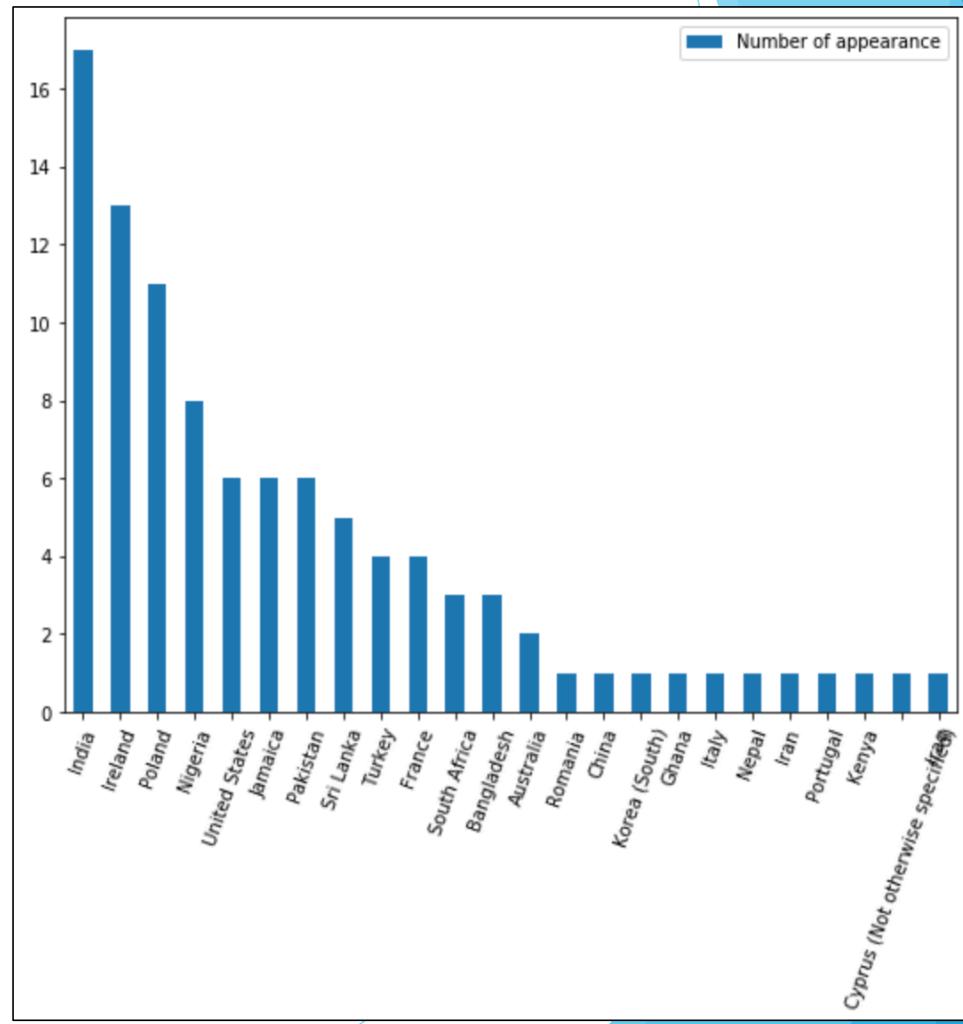
- The number of Indian restaurants
- The number of likes
- The rating

Data gathering and cleaning



Data gathering and cleaning

- India: Highest number of appearance
- Pakistan: 7th
- Sri Lanka: 8th



Data filtering

Conditions applied to boroughs:

- Crime rate < 75th percentile
- Density from India/Pakistan/Sri Lanka > 2 people per hectare
- Gross pay > £30,000
- Median house price < £500,000

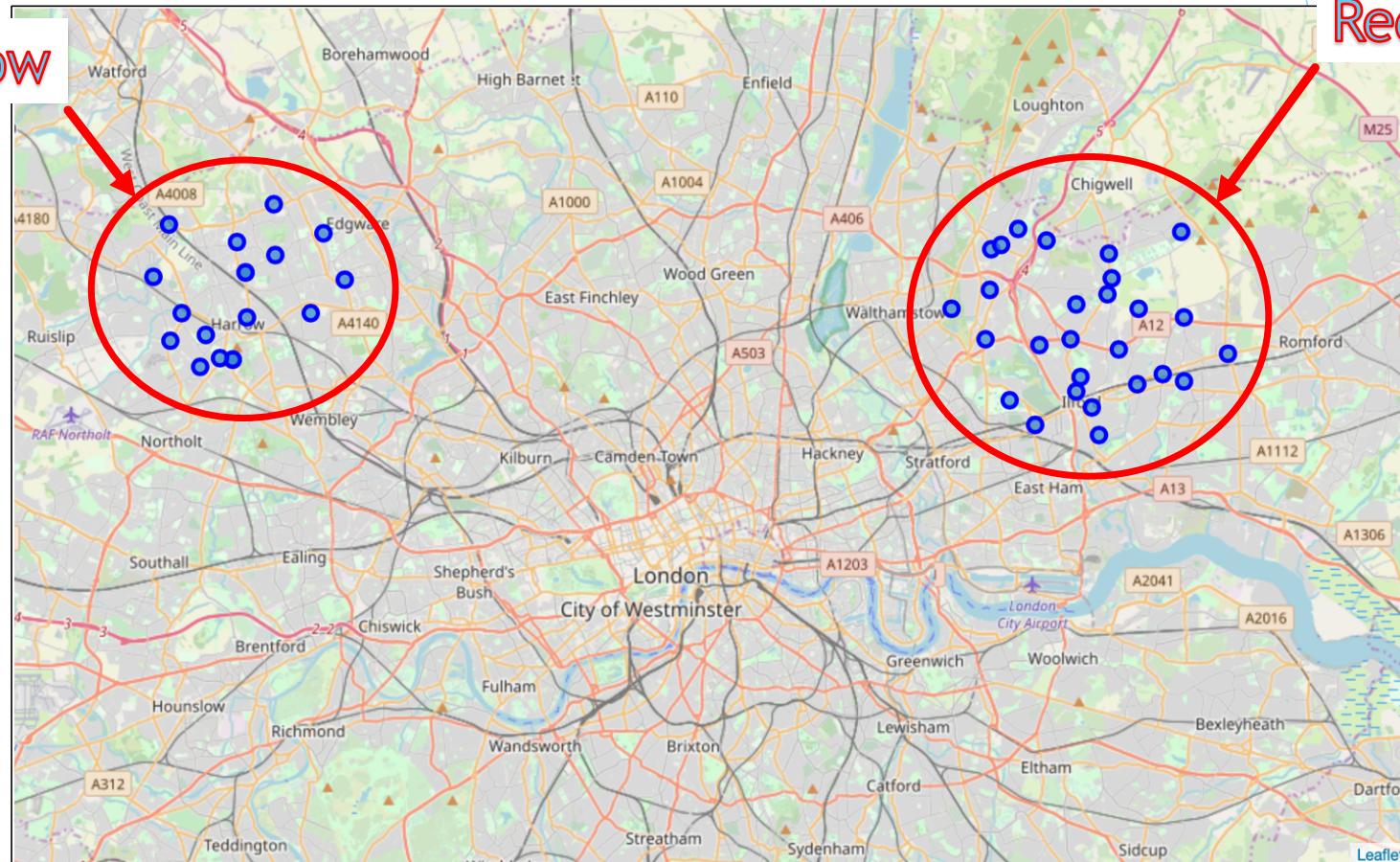
We left with only 2 boroughs: **Harrow and Redbridge**

Borough	Population density (per hectare) 2016	Gross Annual Pay, (2015)	Crime rates per thousand population 2014/15	Median House Price, 2014	Population density from India/Pakistan/Sri Lanka
Harrow	49.8	32529.0	50.4	370000.0	6.6234
Redbridge	53.3	35665.0	69.7	301500.0	8.2615

Map of neighborhoods

Harrow

Redbridge



There is a total of 42 neighborhoods

Foursquare data

Data retrieved from Foursquare for each neighborhood:

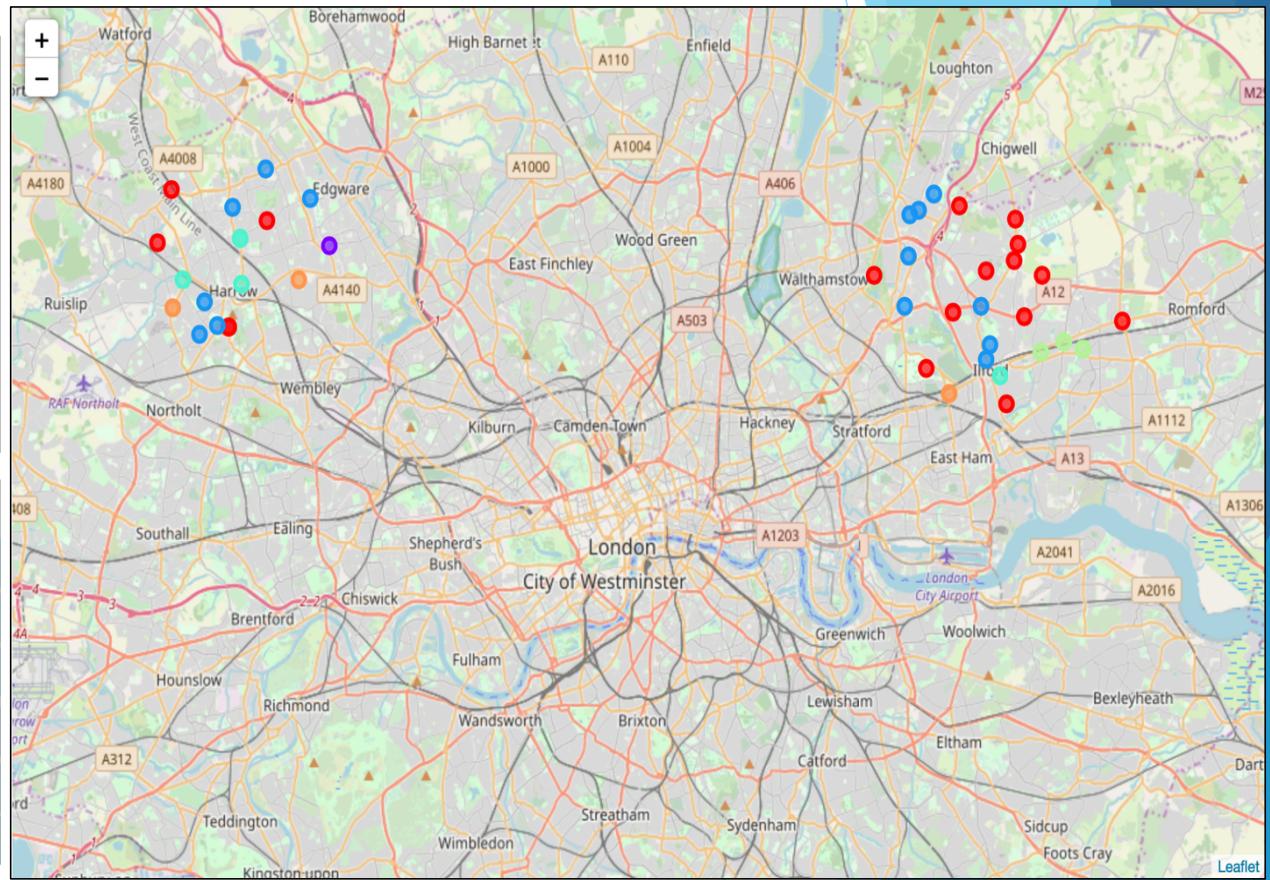
- The number of Indian restaurants
- The highest restaurants' ratings
- The total number of likes

Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Number of restaurants	Max rating	Total number of likes
Aldborough Hatch	51.585525	0.098766	1	0.0	0
Aldersbrook	51.551474	0.048841	13	7.8	64
Barkingside	51.589925	0.083874	4	0.0	4
Belmont	51.601249	-0.319275	6	0.0	3
Canons Park	51.607674	-0.296236	5	7.1	16
Chadwell Heath	51.572228	0.141995	4	0.0	1
Clayhall	51.586974	0.068774	5	0.0	2

Clustering

Cluster number	Number of restaurants		Max rating		Total number of likes	
	max	min	max	min	max	min
0	7	1	0.0	0.0	7	0
1	20	16	7.6	7.0	39	27
2	12	1	8.4	5.8	27	4
3	20	20	8.1	8.1	165	165
4	16	10	0.0	0.0	9	3
5	23	13	7.8	7.6	81	64

Cluster label	Color on map	Description
0	Red	Low competition & poor level of interest
1	Purple	High competition & good level of interest
2	Blue	Medium competition & good level of interest
3	Turquoise	Very High competition & great level of interest
4	Green	Medium competition & poor level of interest
5	Orange	High competition & great level of interest



We would recommend neighborhoods in cluster 2 (blue on map)

Conclusion

- Based on the requirements from the stakeholder and the demographic data of London, only 2 boroughs are qualified for the new location of our Indian restaurant. Harrow and Redbridge.
- There are a total of 42 neighborhoods in Harrow and Redbridge. Using data from Foursquare, the 42 neighborhoods can be clustered

Cluster label	Color on map	Description
0	Red	Low competition & poor level of interest
1	Purple	High competition & good level of interest
2	Blue	Medium competition & good level of interest
3	Turquoise	Very High competition & great level of interest
4	Green	Medium competition & poor level of interest
5	Orange	High competition & great level of interest

- We recommend the stakeholder to open his new restaurant in one of the neighborhoods in cluster 2, which includes a total of 14 neighborhoods.
- The stakeholder is free to make a different choice, depending on what competition level of population interest level they are looking for.