



Open a Restaurant In London

A Machine Learning approached report

Business Problem



London is a huge city, not everyone knows London well before they go there. It is risky to move to a place you barely known.



This report addressed this problem and provide an investment guide to people who want to open a restaurant in London.

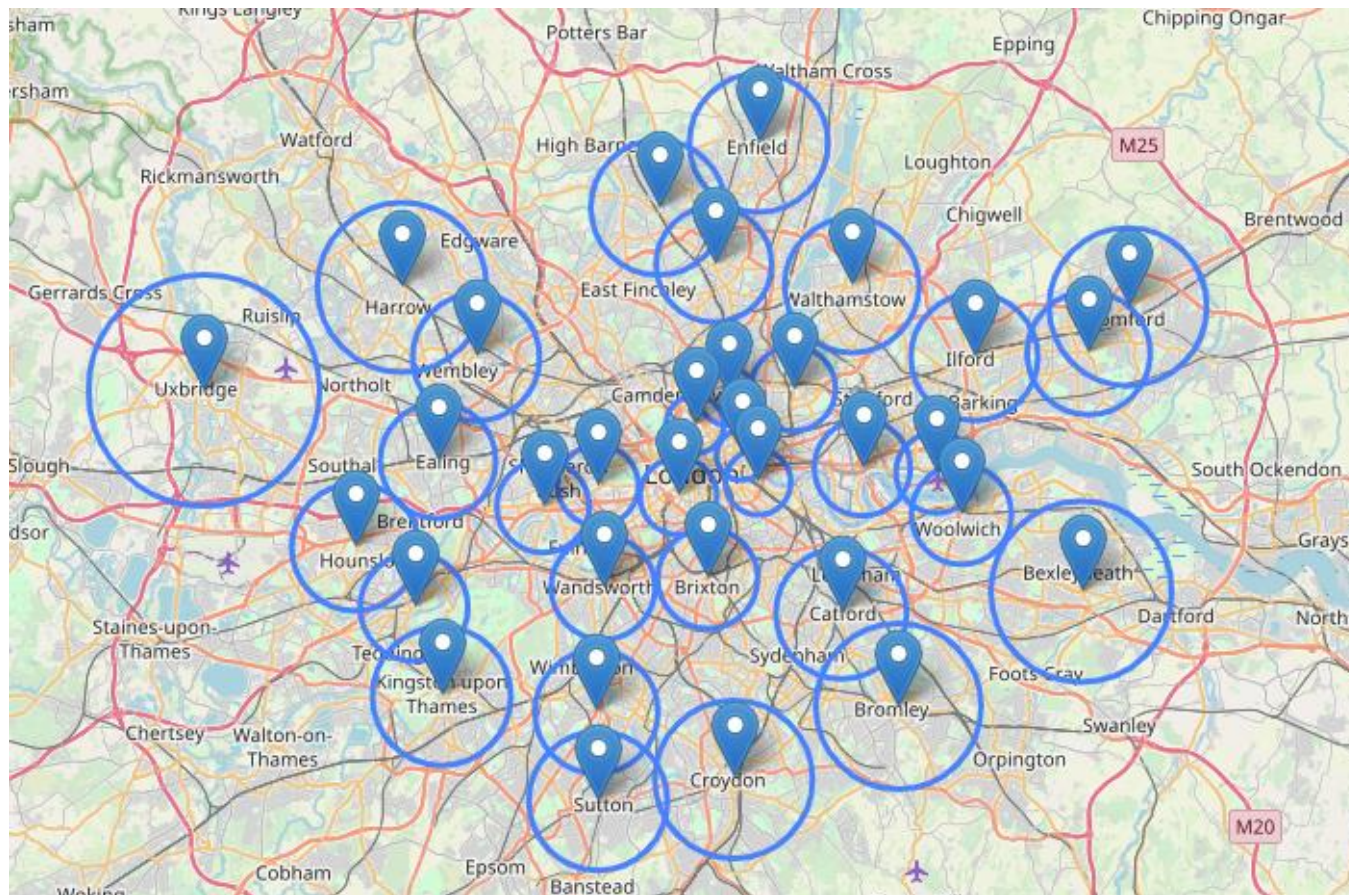


Data acquisition

- List of boroughs can be obtained from Wikipedia. https://en.wikipedia.org/wiki/List_of_London_boroughs
- By using foursquare API, list of recommended venues can be fetched base on above coordinates. <https://api.foursquare.com/v2/venues/explore>
- House sale price data on Land registry of UK, I will use data from last two years <https://use-land-property-data.landregistry.gov.uk/>
- In total 201876 sales record are analyzed
- 3137 records of venues information are used to categories London boroughs

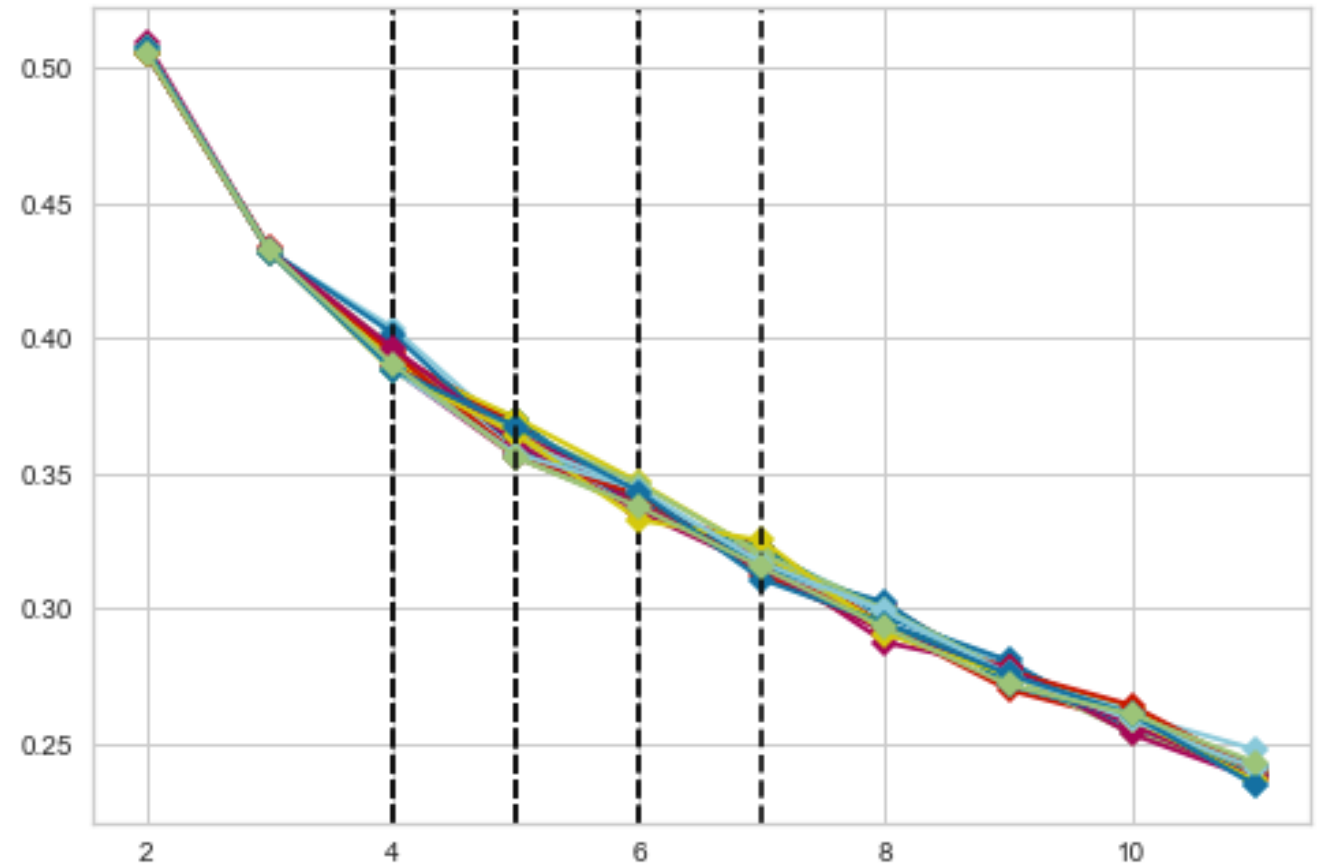
Foursquare API

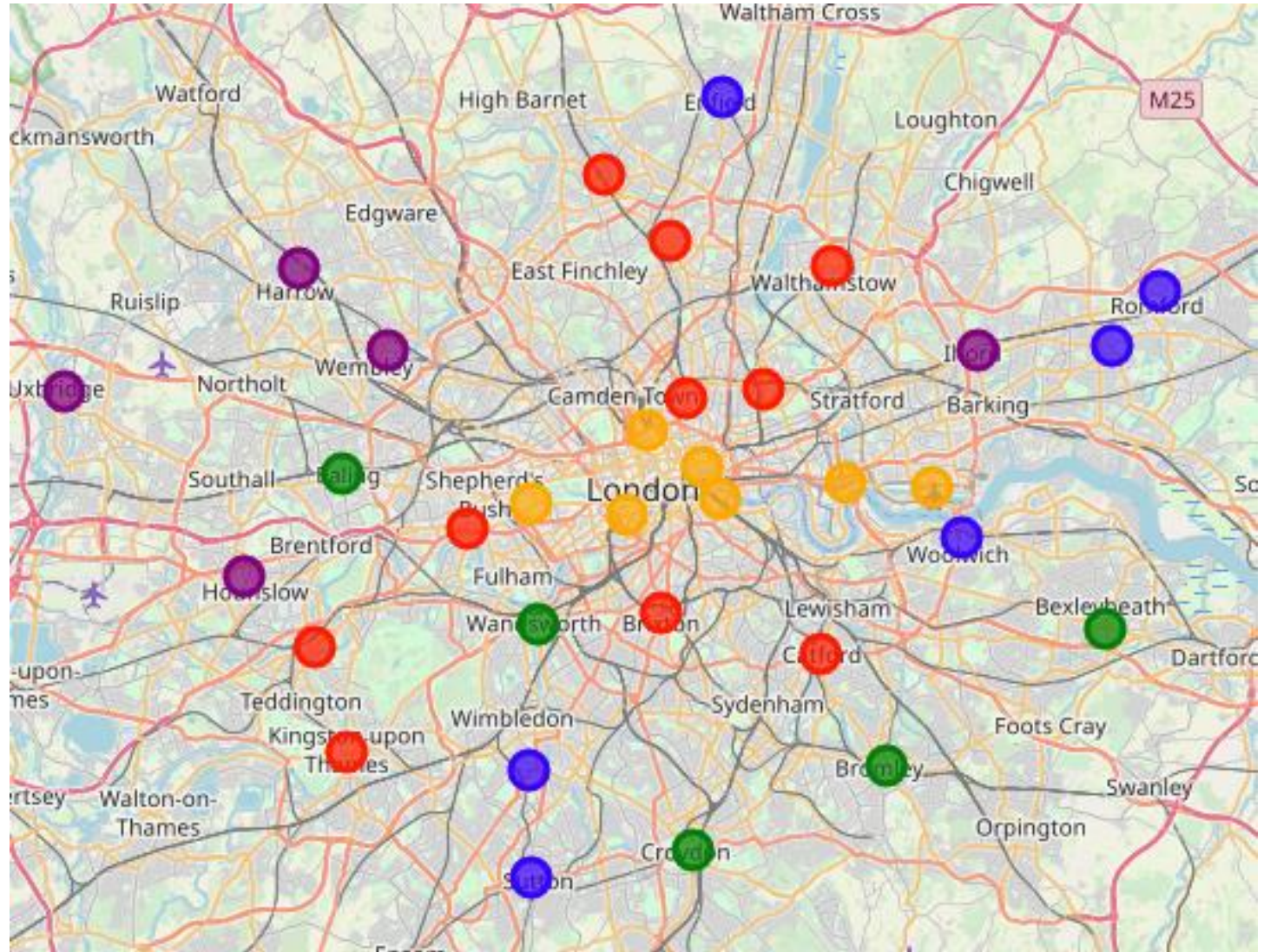
- Optimized radius try to get accuracy data from foursquare API to better represent a borough
- Use the average distance of closest 3 boroughs



K-Mean Elbow Method

- 20 iterations return 17 valid results
- [6, 4, 5, 5, 4, 5, 6, 5, 4, 4, 4, 4, 5, 5, 5, 7, 5]
- Choose 5 as average value





Cluster 0

- Those areas have good night lives because of those pubs, Cafe, and coffee shops.
- If you want to open a restaurant, go to HAMMERSMITH AND FULHAM, and try to avoid pizza, French restaurant, Japanese restaurant and Turkish restaurant.

	borough	latitude	longitude	price	radius	cluster	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
11	HAMMERSMITH AND FULHAM	51.4927	-0.2339	740000.0000	2263	0	Pub	Café	Coffee Shop	Gastropub	Pizza Place	Park	French Restaurant	H
25	RICHMOND UPON THAMES	51.4479	-0.3260	623945.0000	2682	0	Pub	Café	Coffee Shop	Park	Rugby Stadium	Italian Restaurant	Garden	H
17	ISLINGTON	51.5416	-0.1022	587000.0000	1400	0	Pub	Coffee Shop	Bakery	Gastropub	Cocktail Bar	Park	Café	Mediterranean Restaurant
10	HACKNEY	51.5450	-0.0553	557087.5000	2062	0	Pub	Coffee Shop	Bakery	Café	Brewery	Bookstore	Pizza Place	Yoga Studio
20	LAMBETH	51.4607	-0.1163	520000.0000	2484	0	Pub	Coffee Shop	Market	Café	Restaurant	Brewery	Japanese Restaurant	Pizza Place
1	BARNET	51.6252	-0.1517	500000.0000	3291	0	Café	Pub	Coffee Shop	Park	Turkish Restaurant	Italian Restaurant	Supermarket	Grocery Store
12	HARINGEY	51.6000	-0.1119	500000.0000	2807	0	Café	Turkish Restaurant	Pub	Coffee Shop	Park	Bakery	Greek Restaurant	Mediterranean Restaurant
19	KINGSTON UPON THAMES	51.4085	-0.3064	480000.0000	3472	0	Pub	Café	Garden	Park	Hotel	Gastropub	Coffee Shop	Gym / Fitness Centre
29	WALTHAM FOREST	51.5908	-0.0134	440000.0000	3298	0	Pub	Café	Coffee Shop	Pizza Place	Park	Restaurant	Turkish Restaurant	Brewery
21	LEWISHAM	51.4452	-0.0209	409995.0000	3253	0	Pub	Park	Café	Coffee Shop	Gastropub	Italian Restaurant	Indian Restaurant	Gym / Fitness Centre

Cluster 1

- Areas for livings. Because of the supermarkets and grocery stores there. With convenient transportations the difference in average house price are not so big.
- If you want to open a restaurant, try to go to MERTON, and try to avoid Sushi, Indian, and Italian restaurants.

	borough	latitude	longitude	price	radius	cluster	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
22	MERTON	51.4014	-0.1958	464100.0000	3028	1	Park	Supermarket	Pub	Sushi Restaurant	Indian Restaurant	Grocery Store	Fast Food Restaurant	Rest
9	GREENWICH	51.4892	0.0648	430000.0000	2483	1	Pub	Park	Grocery Store	Hotel	Gym / Fitness Center	Café	Coffee Shop	E
8	ENFIELD	51.6538	-0.0799	385550.0000	3445	1	Supermarket	Pub	Turkish Restaurant	Coffee Shop	Grocery Store	Train Station	Gym / Fitness Center	
27	SUTTON	51.3618	-0.1945	380000.0000	3432	1	Pub	Grocery Store	Supermarket	Coffee Shop	Park	Hotel	Italian Restaurant	F C
14	HAVERING	51.5812	0.1837	360000.0000	3918	1	Grocery Store	Coffee Shop	Supermarket	Pub	Bakery	Italian Restaurant	Fast Food Restaurant	
0	BARKING AND DAGENHAM	51.5607	0.1557	309000.0000	3039	1	Grocery Store	Coffee Shop	Park	Supermarket	Pub	Fast Food Restaurant	Shopping Mall	Furn

Cluster 2

- Places mixed with nature and inhabitants.
- Open a restaurant in WANDSWORTH, and avoid pizza, burger, French and Thai restaurant.

	borough	latitude	longitude	price	radius	cluster	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
30	WANDSWORTH	51.4567	-0.1910	610000.0000	2541	2	Coffee Shop	Pub	Park	Pizza Place	Café	Supermarket	Bakery	French Restaurant
7	EALING	51.5130	-0.3089	460000.0000	2839	2	Park	Pub	Coffee Shop	Grocery Store	Hotel	Café	Burger Joint	Pizza Place
4	BROMLEY	51.4039	0.0198	430000.0000	4055	2	Pub	Coffee Shop	Park	Gym / Fitness Center	Clothing Store	Grocery Store	Pizza Place	Indian Restaurant
6	CROYDON	51.3714	-0.0977	372500.0000	3894	2	Coffee Shop	Pub	Grocery Store	Pizza Place	Café	Park	Indian Restaurant	Bookstore
2	BEXLEY	51.4549	0.1505	348620.0000	4513	2	Pub	Grocery Store	Supermarket	Italian Restaurant	Furniture / Home Store	Park	Coffee Shop	Clothing Store

Cluster 3

- Those areas have lots of Indian restaurant. The house price is quite average here.
- Go to BRENT for slightly higher price. Avoid open another Indian restaurant in those areas.

	borough	latitude	longitude	price	radius	cluster	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
3	BRENT	51.5588	-0.2817	470000.0000	3089	3	Coffee Shop	Indian Restaurant	Sandwich Place	Hotel	Clothing Store	Gym / Fitness Center	Grocery Store	Spa
13	HARROW	51.5898	-0.3346	454500.0000	4134	3	Indian Restaurant	Coffee Shop	Pub	Gym / Fitness Center	Supermarket	Café	Park	It Resta
24	REDBRIDGE	51.5590	0.0741	415000.0000	3116	3	Indian Restaurant	Park	Supermarket	Fast Food Restaurant	Clothing Store	Coffee Shop	Portuguese Restaurant	Sanc F
15	HILLINGDON	51.5441	-0.4760	407000.0000	5652	3	Pub	Indian Restaurant	Coffee Shop	Supermarket	Gym / Fitness Center	Golf Course	Park	
16	HOUNSLOW	51.4746	-0.3680	400000.0000	3099	3	Pub	Coffee Shop	Indian Restaurant	Clothing Store	Grocery Store	Hotel	Bakery	Superm

Cluster 4

- Areas with highest house price, have lots of hotels and coffee shop, good places for tourism and business.
- Go to KENSINGTON AND CHELSEA if possible and avoid opening Italian and Indian restaurant.

	borough	latitude	longitude	price	radius	cluster	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
18	KENSINGTON AND CHELSEA	51.5020	-0.1947	1270000.0000	1997	4	Restaurant	Hotel	Gym / Fitness Center	Italian Restaurant	Bakery	Indian Restaurant	Pub	Clothing Store
31	CITY OF WESTMINSTER	51.4973	-0.1372	970000.0000	1891	4	Hotel	Theater	Plaza	Art Museum	Garden	Cocktail Bar	Japanese Restaurant	Indian Restaurant
32	CITY OF LONDON	51.5155	-0.0922	923080.0000	1218	4	Gym / Fitness Center	Coffee Shop	Hotel	Steakhouse	Garden	Cocktail Bar	Vietnamese Restaurant	Sale Place
5	CAMDEN	51.5290	-0.1255	770000.0000	1419	4	Coffee Shop	Café	Hotel	Pizza Place	Bookstore	History Museum	Beer Bar	Burger Joint
26	SOUTHWARK	51.5035	-0.0804	550000.0000	1632	4	Coffee Shop	Hotel	Cocktail Bar	Italian Restaurant	Bakery	Garden	Scenic Lookout	Restaurant
28	TOWER HAMLETS	51.5099	-0.0059	475000.0000	2348	4	Coffee Shop	Hotel	Burger Joint	Gym / Fitness Center	Pub	Bar	Park	Plaza
23	NEWHAM	51.5077	0.0469	402500.0000	2014	4	Hotel	Coffee Shop	Café	Gym / Fitness Center	Sandwich Place	Pub	Supermarket	Grocery Store

Conclusion

- Built a generic categorization model for cities, to generate an overview of a city
- Accuracy of the model can be improved by
 - Get more precise data to indicate the purchase power of boroughs
 - Get more precise data to describe the borough better
 - ❑ Obtain data by more accurate geo shape
 - ❑ Obtain more venues to be used in categorization process