

Coursera Capstone Project Final Report

Finding where to best open a restaurant in London

Thao Dao

2020 – 01 – 20

I. Introduction

London is one of the most expensive and busy cities in the world. This exercise helps to explore potentials for opening a dining business in this city. Anyone who is interested in this topic can find this search helpful.

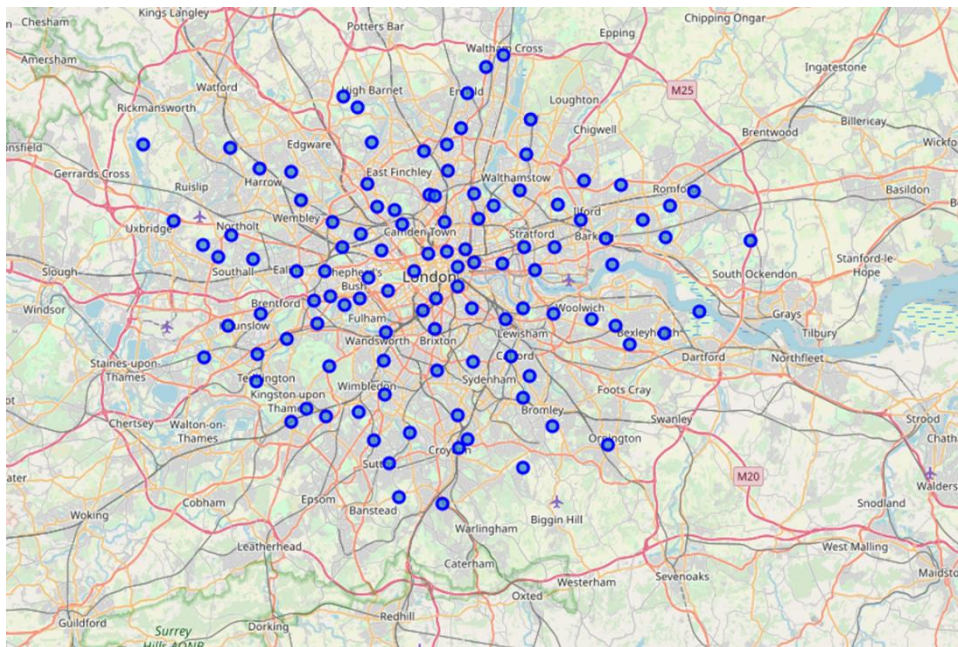
II. Data

There are two data sources for this exercise

- London region post code data
(<https://www.doogal.co.uk/UKPostcodesCSV.ashx?region=E12000007>)
 - This data consists of the zones and district data in London with specific coordinates corresponding with the postal codes
 - The data is grouped by zones and districts where coordinates are taken as average values
- Foursquare data

III. Methodology

- Geographical Exploration : Folium is used to generate the map of London with markers on the districts



- Foursquare used to find nearby venues and one hot encoding to analyse the data

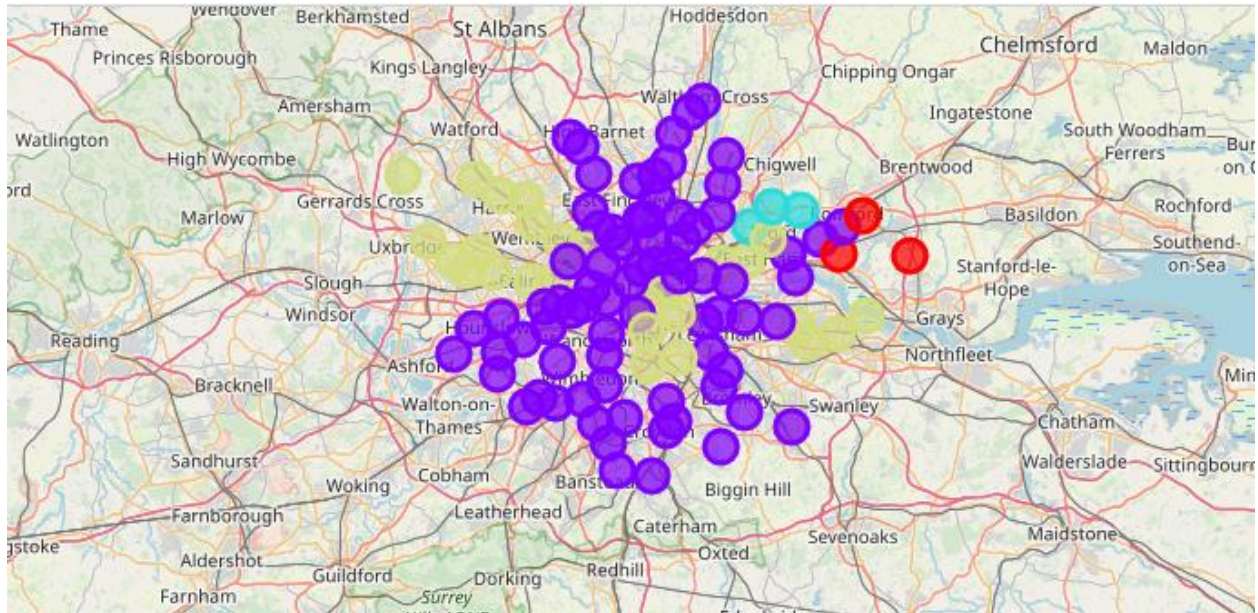
Foursquare

	Neighbourhood	Neighbourhood Latitude	Neighbourhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Camden	51.524655	-0.129067	Bloomsbury Coffee House	51.525282	-0.126495	Café
1	Camden	51.524655	-0.129067	Lever & Bloom Coffee	51.523334	-0.131076	Coffee Shop
2	Camden	51.524655	-0.129067	Tavistock Square	51.525006	-0.129068	Park
3	Camden	51.524655	-0.129067	Gordon Square	51.524098	-0.130763	Park
4	Camden	51.524655	-0.129067	Gay's The Word	51.525386	-0.125369	Bookstore

One hot encodin

	Neighbourhood	Accessories Store	Adult Boutique	Afghan Restaurant	African Restaurant	American Restaurant	Antique Shop	Argentinian Restaurant	Art Gallery	Art Museum	Art ...	Vietnamese Restaurant	Warehouse Store	Watch Shop	Waterfront
0	Barking and Dagenham	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.066667	0.000000	0.000000
1	Barnet	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.055556	0.000000	0.000000	...	0.000000	0.000000	0.000000	0.000000
2	Bexley	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.000000	0.000000	0.000000
3	Brent	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.000000	0.000000	0.000000
4	Bromley	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.000000	0.000000	0.000000
5	Camden	0.000000	0.000000	0.00	0.000000	0.005263	0.000000	0.005263	0.005263	0.005263	...	0.000000	0.000000	0.000000	0.000000

- Modelling using K-means clustering: Different values of kclusters were set to find the best kclusters value based on visualization of data (most effective and quickest to sense check the clusters). Afterwards, each cluster was looked at to see the most common venues.



IV. Results and discussions

k = 4 was found to be the best value for the analysis. The results of 4 clusters are shown below.

[92]:

	London zone	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
85	5	Chinese Restaurant	Bar	Indian Restaurant	Grocery Store	Event Service	Park	Fish & Chips Shop	Farmers Market	Fast Food Restaurant	Field
100	6	Chinese Restaurant	Bar	Indian Restaurant	Grocery Store	Event Service	Park	Fish & Chips Shop	Farmers Market	Fast Food Restaurant	Field
108	7	Chinese Restaurant	Bar	Indian Restaurant	Grocery Store	Event Service	Park	Fish & Chips Shop	Farmers Market	Fast Food Restaurant	Field

Cluster 0 is located further away to East London. If Chinese food is popular, it might be possible that a high population of Chinese community is located there. Seems like specific cuisines are in favour

Cluster 0 is located further away to East London. If Chinese food is popular, it might be possible that a high population of Chinese community is located there. Seems like specific cuisines are in favour.

[96]:

	London zone	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	1	Café	Coffee Shop	Pub	Bakery	Italian Restaurant	Hotel	Bookstore	Pizza Place	Park	Museum
1	1	Coffee Shop	Gym / Fitness Center	Italian Restaurant	Steakhouse	Wine Bar	Hotel	Seafood Restaurant	French Restaurant	Scenic Lookout	Sushi Restaurant
2	1	Coffee Shop	Café	Pub	Bar	Cocktail Bar	Hotel	Vietnamese Restaurant	Yoga Studio	Italian Restaurant	Burger Joint
3	1	Pub	Café	Coffee Shop	Sandwich Place	Park	Pizza Place	Bar	Italian Restaurant	Wine Shop	Gym / Fitness Center
4	1	Hotel	Science Museum	Pub	Bakery	Exhibit	Garden	Coffee Shop	Italian Restaurant	Pizza Place	Café
7	1	Coffee Shop	Hotel	Indian Restaurant	Pub	Clothing Store	Pizza Place	Café	Italian Restaurant	Flea Market	Canal Lock
8	1	Café	Pub	Coffee Shop	Grocery Store	Bus Stop	Gym / Fitness Center	Hotel	Platform	Park	Cocktail Bar
9	1	Clothing Store	Café	Hotel	Boutique	Italian Restaurant	Art Gallery	Pub	Coffee Shop	Lounge	Japanese Restaurant
10	2	Grocery Store	Golf Course	Indian Restaurant	Athletics & Sports	Bus Stop	Stables	Market	Tennis Court	Beer Garden	Coffee Shop
12	2	Café	Coffee Shop	Pub	Bakery	Italian Restaurant	Hotel	Bookstore	Pizza Place	Park	Museum
14	2	Boat or Ferry	Pub	Grocery Store	Garden	Bus Stop	Café	Bakery	Market	Burger Joint	History Museum

Interesting to see cluster 1 is overwhelmed with cafe, leisure and restaurants. It is randomly distributed all over 7 zones in London. Here are more like shopping malls or busy streets where coffee shops are very popular

Interesting to see cluster 1 is overwhelmed with cafe, leisure and restaurants. It is randomly distributed all over 7 zones in London. Here are more like shopping malls or busy streets where coffee shops are very popular

[94]:

	London zone	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
49	3	Bowling Alley	Middle Eastern Restaurant	Pub	Fast Food Restaurant	Hookah Bar	Pizza Place	Farmers Market	Field	Fish & Chips Shop	Flea Market
73	4	Bowling Alley	Middle Eastern Restaurant	Pub	Fast Food Restaurant	Hookah Bar	Pizza Place	Farmers Market	Field	Fish & Chips Shop	Flea Market
89	5	Bowling Alley	Middle Eastern Restaurant	Pub	Fast Food Restaurant	Hookah Bar	Pizza Place	Farmers Market	Field	Fish & Chips Shop	Flea Market

Cluster 2 is mostly for Middle East population, bowling alley seems to be the local activity. Fastfood or Pub are more popular

Cluster 2 is mostly for Middle East population, bowling alley seems to be the local activity. Fastfood or Pub are more popular

[95]:

	London zone	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
5	1	Pub	Hotel	Restaurant	Gym / Fitness Center	Park	Portuguese Restaurant	Café	Art Gallery	Arts & Crafts Store	Nightclub
6	1	Pub	Park	Fast Food Restaurant	Café	Garden	Sandwich Place	Argentinian Restaurant	Italian Restaurant	Bus Stop	Cricket Ground
11	2	Park	Fish & Chips Shop	Yoga Studio	Bar	Pub	Department Store	Fast Food Restaurant	Café	Deli / Bodega	Metro Station
13	2	Pub	Café	Park	Hotel	Gym / Fitness Center	Mini Golf	Chinese Restaurant	Gas Station	Train Station	Scenic Lookout
21	2	Pub	Hotel	Restaurant	Gym / Fitness Center	Park	Portuguese Restaurant	Café	Art Gallery	Arts & Crafts Store	Nightclub
23	2	Bus Stop	Café	Fast Food Restaurant	Historic Site	Pub	Supermarket	Boutique	Park	Gym / Fitness Center	Waterfront
25	2	Pub	Park	Fast Food Restaurant	Café	Garden	Sandwich Place	Argentinian Restaurant	Italian Restaurant	Bus Stop	Cricket Ground
32	3	Park	Fish & Chips Shop	Yoga Studio	Bar	Pub	Department Store	Fast Food Restaurant	Café	Deli / Bodega	Metro Station
36	3	Pub	Café	Park	Hotel	Gym / Fitness Center	Mini Golf	Chinese Restaurant	Gas Station	Train Station	Scenic Lookout
45	3	Pub	Hotel	Restaurant	Gym / Fitness Center	Park	Portuguese Restaurant	Café	Art Gallery	Arts & Crafts Store	Nightclub
48	3	Bus Stop	Café	Fast Food Restaurant	Historic Site	Pub	Supermarket	Boutique	Park	Gym / Fitness Center	Waterfront

Cluster 3 is quite similar to cluster 1 but with more pub - more "localness" in it. It goes up to the northwest area of London.

Cluster 3 is quite similar to cluster 1 but with more pub - more "localness" in it. It goes up to the northwest area of London.

V. Conclusions

It does look like the algorithm is not splitting well between class 1 and class 3 but it might be because this is the nature of London. London is a service city where restaurants are everywhere. Depending on where the location is, what type of cuisine or whether this is a coffee shop, pub, etc. one can pick the right combination to open the business