

Analysis of Neighborhood Restaurants in London

Ebin Paul

1. Introduction

In this project, we'll try to solve a problem which usually happens to people who are visiting a new city which is, what are the good affordable restaurants in a given neighborhood/town that serve good and affordable cuisines of one's liking? So, the scope of this project will be to provide a list of good restaurants which provide the cuisine of their choice also it should be hygienic and should have good ratings as well.

Since London is one of the many multicultural cities around the world, we will be comparing the neighborhoods of London and cumulating a list of restaurants based on the cuisine they serve, and their ratings. For this to work we'll rely on data collected from different sources and various types of visualizations including plotting the areas on the map.

2. Data

Data Sources

For this project we'll be gathering public data from following sources:

1. **Wikipedia:** Here we can get the data related to boroughs and the locations within those boroughs. Since this data is based on the locations under each borough in London, we can group these areas based on the boroughs they belong to. Click [here](#) to visit the Wikipedia page.
2. **Foursquare:** Using Foursquare API we can obtain longitudes and latitudes of all the boroughs and using that data we can plot the locations and popular venues in London. Also using this API, we can also get the details of all the restaurants that serve different types of cuisines from around the world.
3. **Data World:** Using Data World we can get the detailed information about the venues which we'll get from Foursquare API like for example ratings.

3. Methodology

Web scraping method was used to obtain the borough and location data. The data from Wikipedia contains postal code for each location, with the postal code data we can obtain the coordinates for the location.

Once coordinates are obtained then using that data, we can obtain data regarding the top venues in a given location. Now the obtained data has to be cleaned. First, columns that have the dial code and OS grid ref data are removed and then locations are segregated based on one borough which in this case is Camden. Below is the list of locations/neighborhoods in Camden:

	Location	Boroughs	Post Town	Postcode
39	Belsize Park	Camden	LONDON	NW3
54	Bloomsbury	Camden	LONDON	WC1
76	Camden Town	Camden	LONDON	NW1
86	Chalk Farm	Camden	LONDON	NW1
174	Fitzrovia	Camden	LONDON	W1
182	Frognal	Camden	LONDON	NW3
192	Gospel Oak	Camden	LONDON	NW5, NW3
210	Hampstead	Camden	LONDON	NW3
239	Highgate	Camden	LONDON	N6
242	Holborn	Camden	LONDON	WC1, WC2
263	Kentish Town	Camden	LONDON	NW5
375	Primrose Hill	Camden	LONDON	NW1, NW3, NW8
411	Somerstown	Camden	LONDON	NW1
431	St Giles	Camden	LONDON	WC2
436	St Pancras	Camden	LONDON	WC1
453	Swiss Cottage	Camden	LONDON	NW3
505	West Hampstead	Camden	LONDON	NW6

Once the data frame with locations in Camden is created then, we'll get the data which contains popular venues in Camden using Foursquare API. The problem with this data is that it will data which contains the venues which are not restaurants as well. Once the data is filtered for restaurants the data frame looks like this:

	Location	Venue	Venue ID	Venue Latitude	Venue Longitude	Venue Category
1	Belsize Park	Patara	574ad2b5498e49a138716f26	51.555497	-0.176086	Thai Restaurant
4	Belsize Park	28 Church Row	579b97f6498e51784c2028e0	51.555610	-0.178724	Tapas Restaurant
9	Belsize Park	Jin Kichi 人吉 (Jin Kichi)	4b0e428af964a5206d5623e3	51.557211	-0.178370	Japanese Restaurant
18	Belsize Park	Paradise	4baa6ff3f964a520106a3ae3	51.555476	-0.166312	Indian Restaurant
21	Belsize Park	Spielburger	5377c796498e21d3bd43dbc0	51.556336	-0.178640	American Restaurant
...
804	West Hampstead	Ruchi	4bc6045c42419521116f031d	51.543091	-0.195521	Indian Restaurant
811	West Hampstead	Vijay India	4c13af69a9c220a1f121569d	51.540883	-0.201734	Indian Restaurant
812	West Hampstead	Speedy Noodle	4bbf963074a9a59345b6cef6	51.542741	-0.198992	Asian Restaurant
813	West Hampstead	McDonald's	4b7800e8f964a520bbb12ee3	51.539607	-0.195667	Fast Food Restaurant
818	West Hampstead	Chasing Dragon	4d45cbf71b62b1f740c400e3	51.544509	-0.201773	Chinese Restaurant

Now from the data obtained from Data World, we'll add the data like Restaurant Rating, Hygiene Score and Structural score. After add these data the final data frame looks like this:

	Location	Venue	Venue ID	Venue Latitude	Venue Longitude	Venue Category	RestaurantRating	HygieneScore	StructuralScore
4	Belsize Park	28 Church Row	579b97f6498e51784c2028e0	51.555610	-0.178724	Tapas Restaurant	4	10	0
21	Belsize Park	Spielburger	5377c796498e21d3bd43dbc0	51.556336	-0.178640	American Restaurant	5	0	0
25	Belsize Park	Dim T	4b52fa6ff964a520d38b27e3	51.555292	-0.178349	Dim Sum Restaurant	4	5	10
38	Belsize Park	Fish Cafe	5309e88811d2e2fbd759bcd8	51.555995	-0.177251	English Restaurant	4	10	5
43	Belsize Park	La Cage Imaginaire	4c4f4e9c92b6a593294f5a71	51.556736	-0.177311	French Restaurant	5	0	5
...
776	Swiss Cottage	La Cage Imaginaire	4c4f4e9c92b6a593294f5a71	51.556736	-0.177311	French Restaurant	5	0	5
794	West Hampstead	Bombay Flame	4ba8be95f964a5209eeb39e3	51.542676	-0.195202	Indian Restaurant	3	5	10
796	West Hampstead	German Doner Kebab	58a8b2229168024c86918f9a	51.544297	-0.200867	Doner Restaurant	3	10	10
803	West Hampstead	Nando's	4c1e5386fcf8c9b66e10ad0b	51.544655	-0.200661	Portuguese Restaurant	5	5	5
813	West Hampstead	McDonald's	4b7800e8f964a520bbb12ee3	51.539607	-0.195667	Fast Food Restaurant	5	0	5

Exploring Data

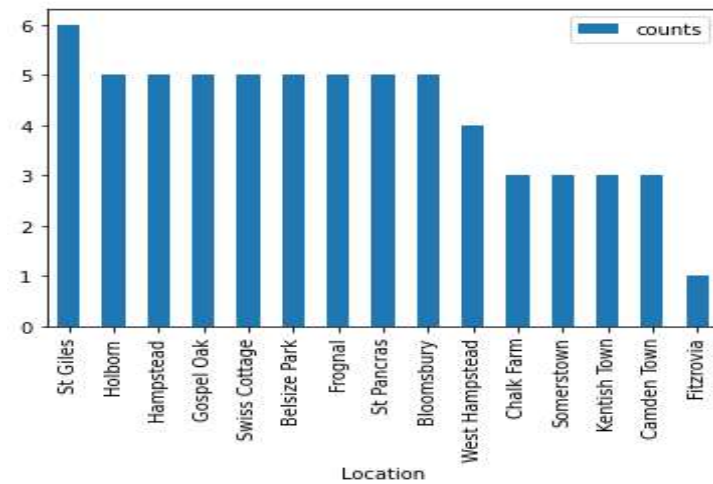
Locations like Barnet, Westminster, Camden, Lewisham etc. are the some of the locations that have more than 15 locations or neighborhoods under them. The venue data obtained from Foursquare API shows that there are a large number of Italian restaurants in Camden.

Italian Restaurant	26
Japanese Restaurant	16
Indian Restaurant	12
French Restaurant	11
Thai Restaurant	10
Tapas Restaurant	9
Sushi Restaurant	9
Argentinian Restaurant	7
English Restaurant	6
Asian Restaurant	6
Fast Food Restaurant	6
Falafel Restaurant	6
American Restaurant	5
Turkish Restaurant	5
Restaurant	5
Dim Sum Restaurant	5
Szechuan Restaurant	3
Chinese Restaurant	3
Brazilian Restaurant	2
Korean Restaurant	2
Seafood Restaurant	2
Ethiopian Restaurant	2
Vegetarian / Vegan Restaurant	2
Modern European Restaurant	1
Kebab Restaurant	1
Middle Eastern Restaurant	1
Vietnamese Restaurant	1
Okonomiyaki Restaurant	1
Lebanese Restaurant	1
Doner Restaurant	1
Portuguese Restaurant	1

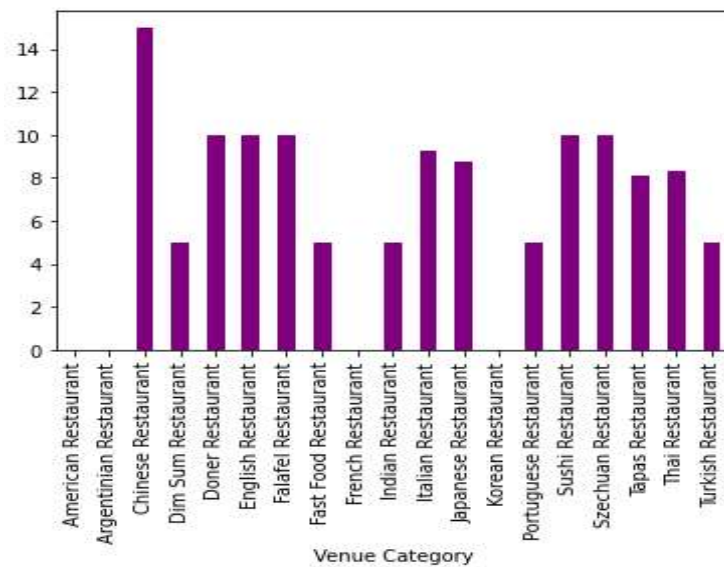
Among the restaurants that serve different types of cuisines, restaurants that serve American, French and Italian cuisines are the ones with high ratings. Meanwhile restaurant serving Korean, Portuguese cuisines are the ones with the least rating.

American Restaurant	5
French Restaurant	5
Italian Restaurant	3
Korean Restaurant	1
Portuguese Restaurant	1
Fast Food Restaurant	1

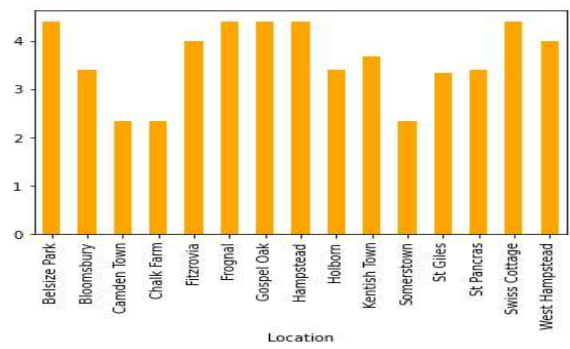
Locations like St. Giles, Holborn, Hampstead etc. are some of the locations with a large number of restaurants.



Restaurant hygiene is one of the many aspects which decides the quality of the food served at the restaurant. Here below are the hygiene score of the restaurants based on the cuisine they serve:



The above graph shows the mean hygiene score of available cuisines in Camden, except from restaurants serving American, Argentinian, French and Korean cuisines all other restaurants have fairly better hygiene score.



The graph above shows the locations along with their average restaurant ratings. The graph suggests that fairly a large number of locations have an average rating more than 3.

After exploring the data further, we noticed that restaurants having higher hygiene and structural score have very low restaurant ratings.

Finally, arranging restaurants based on the ratings they have:

	Location	Venue	Venue ID	Venue Latitude	Venue Longitude	Venue Category	RestaurantRating	HygieneScore	StructuralScore
411	Hampstead	La Cage Imaginaire	4c4f4e9c92b6a593294f5a71	51.556736	-0.177311	French Restaurant	5	0	5
283	Frogna	Spielburger	5377c796498e21d3bd43dbc0	51.556336	-0.178640	American Restaurant	5	0	0
803	West Hampstead	Nando's	4c1e5386fcf8c9b66e10ad0b	51.544655	-0.200661	Portuguese Restaurant	5	5	5
776	Swiss Cottage	La Cage Imaginaire	4c4f4e9c92b6a593294f5a71	51.556736	-0.177311	French Restaurant	5	0	5
754	Swiss Cottage	Spielburger	5377c796498e21d3bd43dbc0	51.556336	-0.178640	American Restaurant	5	0	0
...
153	Chalk Farm	Trufflesecco	57112d67498eebb7f35ec4f	51.534781	-0.139133	Italian Restaurant	2	15	5
721	St Pancras	Chilli Cool	4add9af0f964a520236521e3	51.526703	-0.124751	Szechuan Restaurant	1	10	15
475	Holborn	Chilli Cool	4add9af0f964a520236521e3	51.526703	-0.124751	Szechuan Restaurant	1	10	15
86	Bloomsbury	Chilli Cool	4add9af0f964a520236521e3	51.526703	-0.124751	Szechuan Restaurant	1	10	15
670	St Giles	Canton Element	5273a803498e5eafdcc026d2	51.519931	-0.117408	Chinese Restaurant	0	15	20

4.Result

We were able to;

1. Get the names of each borough in London along with the locations under each borough.
2. Gather the coordinates for each location.
3. Gather the top venues in a borough.
4. Obtain the required data for the restaurants like, name, location, category, rating, hygiene and structural scores.
5. Explore the data regarding the restaurant and plot the required plots.
6. Plot the areas, and restaurants in a map.

5. Limitations

1. The model's accuracy is based on the accuracy of the data obtained.
2. There is a limitation on the number of API calls which restricts the amount of data which we can collect.

6. Conclusion

This model can be used for obtaining the required data regarding venues like restaurants, bars, parks, cafes etc. Foursquare API can be used to get the required details regarding the venues in a given location. Also, this model can be used to recommend restaurants based on the location, type of cuisine, ratings, hygiene, quality of the place etc.