

The Battle of the Neighbourhoods

1. Introduction

1.1 Background

"UK restaurant market facing fastest decline in seven years"

A headline from last yearⁱ prior to the coronavirus.

MCA's UK Restaurant Market Report 2019ⁱⁱ indicated that *"large falls in the sales value and outlet volumes of independent restaurants is the cause of the overall decline of the UK restaurant market. It attributes this to a "perfect storm" of rising costs, over-supply, and weakening consumer demand."*

London's restaurant scene changes week on week, with openings and closures happening on a regular basis; it must be hard to keep up. The hyper-competitiveness of London's restaurant scene make it one of the toughest cities in the world to launch a new venture.

"With business rates up and footfall down, a winning formula is worth its weight in gold and although first-rate food is inevitably the focus, other factors can also affect a restaurant's success. Atmosphere is frequently cited in customer surveys as second only to food in an enjoyable restaurant visit and getting the vibe right is crucial."ⁱⁱⁱ

Due to the coronavirus most businesses have suffered even greater losses. As restrictions lift businesses will be looking for ways to make up for lost time and earnings. Reopening a restaurant once lockdown is over is one thing, but knowing what to put on the menu if you haven't been in contact with a punter in months is another.

"Are there any grounds for hope? A wild optimist might point to some encouraging data about the overperformance of small chains while everyone else loses their shirts; a realist might make coughing noises about small sample sizes and growth from a low base. The queues snaking out of Soho's recently opened Pastaio suggest one genuinely viable route to salvation - concepts may need to follow its lead and amp up the comfort food factor while dialling down prices.

And while home delivery is a source of confidence for some parties (Deliveroo, for instance, recently listed its shares on the stock market) it may well end up a false friend: the increased volume of so-called "dark kitchens" presage a sinister vision of the future, where restaurants don't exist to serve customers onsite at all, but just pump out takeaway meals for us to consume on our sofas. A little far-fetched, perhaps, but with lights going out at a faster rate than many can remember, it can't be too long before whole tranches of the market do indeed go dark, one way or another."^{iv}

ⁱ <https://www.bighospitality.co.uk/Article/2019/09/24/UK-restaurant-market-facing-fastest-decline-in-seven-years-according-to-MCA-Insight>

ⁱⁱ <https://www.mca-insight.com/market-reports/uk-restaurant-market-report-2019/597394.article>

ⁱⁱⁱ <https://www.standard.co.uk/go/london/restaurants/how-londons-top-restaurants-soundtrack-their-spaces-a3737991.html>

^{iv} <https://www.theguardian.com/global/2017/nov/26/who-killed-londons-restaurant-scene>

1.2 Business problem

The task is to identify a new, on trend hospitality business opportunity in a thriving location in London. However, with the country currently under lockdown it is much harder to understand what types of food and drink businesses are popular. A novel method will need to be deployed to analyse the current situation with food businesses in London during the coronavirus.

The project originally planned to use foot traffic data at different times to identify what food venues appear to be trending and where. However, given that the whole country is under lockdown due to the Coronavirus there is no trending data to analyse.

Another flawed idea would be to assume whatever food venues are most common are most in demand/popular. This is not ideal since it would be using data too set in the past (it takes time to build a restaurant) and trends move more quickly. This method would not provide near-real-time visibility of what is trending to make more accurate predictions, ahead of the curve. In other words, market trends and tastes change all the time and whatever shows up using the mode of venue results is only what was popular months prior. **Note:** This is true of this method, regardless of lockdown!

Instead, data that might contribute to determining restaurant improvements might include performance metrics during lockdown; hours, venue likes, volume of recommendations, quality of recommendations, content of recommendations (word densities), as well as clusters of food businesses that remained operational during lockdown. This project aims to predict pandemic proof food enterprises and what the industry might look like after restrictions are lifted.

1.3 Interest

Obviously, any restaurateur or leisure and hospitality entrepreneur/enterprise would be extremely interested in accurate prediction of food trend data for competitive advantage and added business value. These such data could be used to inform new menu creation or concepts for new boutique restaurants or street food vendor pop-ups - proving valuable to food and retail parks, such as Boxpark.

1.4 Desired outcome

The ideal outcome of this notebook would be to:

- Create a dataframe of London districts, postcode centroids, and coordinates
- Identify the top 5 (most common) food venues by cuisine
- Plot all food venues to map
- Use venue Hours Endpoint of FS API to see what venues are still operating during lockdown
- Create word clouds of venue Tips using Foursquare API to identify trending menu items. Since the Foursquare trending feature won't return any results at this time due to the coronavirus lockdown, I will use the Tips Endpoint in the Foursquare API to try and identify patterns in the reviews i.e. what menu items get the most positive mentions (**Note:** This may require the use of sentiment analysis, which is out of scope for this notebook)*

- Use choropleth maps to highlight food vendor densities per London district by different cuisines (optional: choropleth map by density of venues open during lockdown*)
- Use k-means clustering to cluster food venues in London to identify restaurant hotspots and prime locations as suggestions for the client (optional: cluster by venues open during lockdown)
- Map commercial venues that are available to rent using either Zoopla API or Rightmove webscraper*
- List suitable commercial venues for further analysis*

**may be out of scope.*

2. Data Acquisition and Cleaning

2.1 Data sources

To explore the problem, we can use the data listed below:

- Wikipedia page for London Postal Districts^v to get an initial high-level view of what we are working with.
- For cross-referencing postcodes with districts and longitudes and latitudes I will use a combination of the Office for National Statistics^{vi} and London Datastore^{vii}. I also checked "A Guide to ONS Geography Postcode Products"^{viii} to make sure I was using the correct postcode system for statistics (NSPL).
- I found the Second-level Administrative Divisions of the United Kingdom from NYU Spatial Data Repository^{ix}. The .json file^x has coordinates and boundaries of the all the cities of the UK. This will be cleaned and reduced to London where I will use it to create a choropleth map of food vendor densities for different cuisines using the Foursquare API.
- Foursquare API^{xi} will be used to get the most common food venues of London. **Note:** This may be reduced further to City of London and Westminster (or just EC and WC postcodes) to reduce the number of API calls. The FS API's hours and tips Endpoints^{xii} will be used to get venue operating hours (to see which venues are still operating during lockdown) and to get user recommendations, which will be used for word clouds to try and identify trending menu items.

^v https://en.wikipedia.org/wiki/London_postal_district

^{vi} <https://geoportal.statistics.gov.uk/datasets/national-statistics-postcode-lookup-may-2020>

^{vii} <https://data.london.gov.uk/london-area-profiles/>

^{viii} <https://www.ons.gov.uk/methodology/geography/geographicalproducts/postcodeproducts>

^{ix} <https://geo.nyu.edu/catalog/stanford-wj438mh2295>

^x <https://earthworks.stanford.edu/download/file/stanford-wj438mh2295-geojson.json>

^{xi} <https://developer.foursquare.com/>

^{xii} <https://developer.foursquare.com/docs/places-api/endpoints/>

- I will then use either Zoopla API^{xiii} or Rightmove webscraper^{xv} to pull in commercial properties on the market as options for the client.

^{xiii} <https://developer.zoopla.co.uk/>

^{xiv} <https://github.com/AnthonyBloomer/zoopla>

^{xv} https://github.com/toby-p/rightmove_webscraper.py

3. Methodology

3.1 Exploratory Data Analysis

3.1.1 Tranform the data into a pandas dataframe and explore data

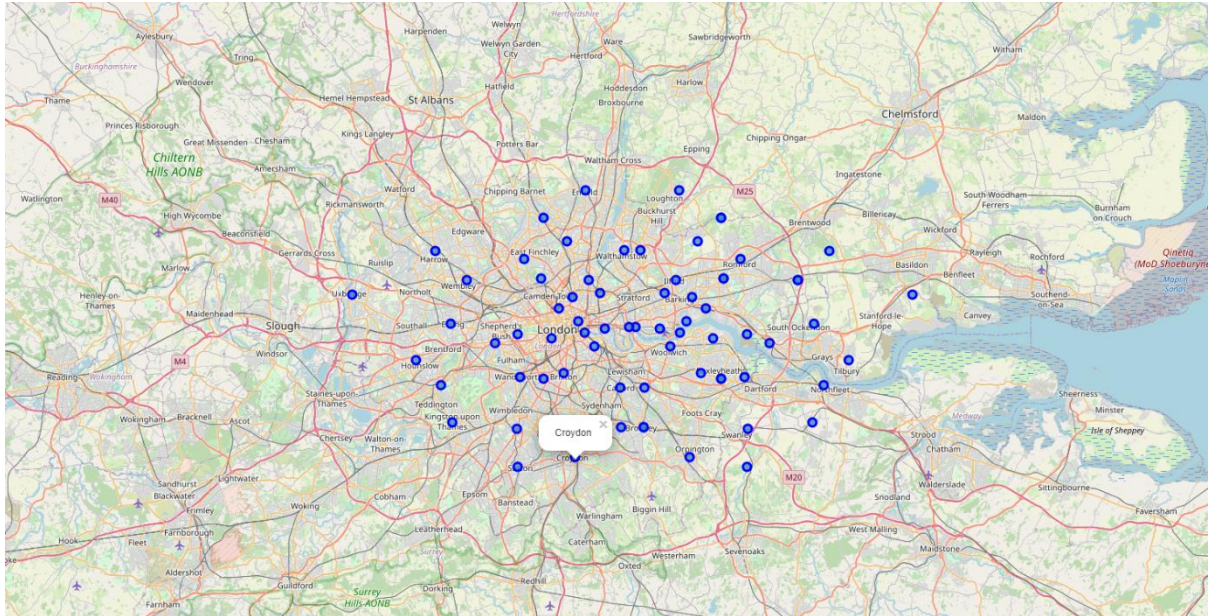
London has a total of 32 boroughs and the City of London and 533 sub-districts/areas/neighborhoods (too many to explore in this study). In order to segment the neighborhoods and explore them, we will use a London Boroughs Dataframe that contains the 32 boroughs as well as the latitude and longitude coordinates of each borough.

Note: Postcodes are irrelevant here, since there are many postcodes within each of these boroughs and we already have the coordinates. If we did not have the coordinates already, we could perform analysis to find the centroids of each cluster of postcodes belonging to each borough.

	Borough	Latitude	Longitude
0	Barking and Dagenham	51.5607	-0.1557
1	Barnet	51.6252	-0.1517
2	Bexley	51.4549	-0.1505
3	Brent	51.5588	-0.2817
4	Bromley	51.4039	-0.0198
5	Camden	51.5290	-0.1255
6	City of London	51.5155	-0.0922
7	Croydon	51.3714	-0.0977
8	Ealing	51.5130	-0.3089
9	Enfield	51.6538	-0.0799
10	Greenwich	51.4892	-0.0648

3.1.2 Transform the data into a pandas dataframe and explore data

First get the location of London and generate the bare map using folium and then add the markers for the Boroughs.



3.1.3 Exploring and Mapping the Venues in the London Boroughs

Now we are going to start utilising the Foursquare API to explore the neighbourhoods and segment them. We make a call to the API to request venues in each of the Boroughs in London along with the venue category.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Barking and Dagenham	51.5607	-0.1557	Parliament Hill	51.559661	-0.159639	Scenic Lookout
1	Barking and Dagenham	51.5607	-0.1557	Parliament Hill Farmers' Market	51.559420	-0.151418	Farmers Market
2	Barking and Dagenham	51.5607	-0.1557	Bistro Laz	51.561333	-0.150528	Mediterranean Restaurant
3	Barking and Dagenham	51.5607	-0.1557	GAIL's Bakery	51.561942	-0.149528	Bakery
4	Barking and Dagenham	51.5607	-0.1557	Parliament Hill Fields Playground	51.556653	-0.157296	Playground
5	Barking and Dagenham	51.5607	-0.1557	Al Parco	51.561356	-0.150680	Italian Restaurant
6	Barking and Dagenham	51.5607	-0.1557	Carob Tree	51.561377	-0.150580	Mediterranean Restaurant
7	Barking and Dagenham	51.5607	-0.1557	Forks & Corks	51.561899	-0.150116	Food
8	Barking and Dagenham	51.5607	-0.1557	Parliament Hill Café	51.559242	-0.152850	Café
9	Barking and Dagenham	51.5607	-0.1557	Kalendar	51.561890	-0.149737	Café
10	Barking and Dagenham	51.5607	-0.1557	Highgate Men's Bathing Pond	51.563351	-0.156817	Lake
11	Barking and Dagenham	51.5607	-0.1557	Parliament Hill Fields Tennis Courts	51.560194	-0.150708	Tennis Court
12	Barking and Dagenham	51.5607	-0.1557	Parliament Hill Athletics Track	51.556979	-0.155811	Track Stadium
13	Barking and Dagenham	51.5607	-0.1557	Parliament Hill Track	51.557033	-0.153230	Trail
14	Barnet	51.6252	-0.1517	The Atrium	51.624726	-0.151933	Café
15	Barnet	51.6252	-0.1517	Made Curtains	51.623485	-0.153565	Home Service
16	Barnet	51.6252	-0.1517	Beaconsfield Road (BF)	51.622827	-0.151466	Bus Stop
17	Barnet	51.6252	-0.1517	Oakleigh Cafe	51.623412	-0.154899	Café
18	Barnet	51.6252	-0.1517	Premium Electrical	51.623175	-0.153420	Business Service
19	Bexley	51.4549	-0.1505	Clapham Common	51.457446	-0.150587	Park
20	Bexley	51.4549	-0.1505	Moxon's Fishmongers	51.453170	-0.147857	Fish Market

'The Battle of the Neighborhoods' Report

Next, we check out all the unique venue categories.

```
array(['Scenic Lookout', 'Farmers Market', 'Mediterranean Restaurant',
      'Bakery', 'Playground', 'Italian Restaurant', 'Food', 'Café',
      'Lake', 'Tennis Court', 'Track Stadium', 'Trail', 'Home Service',
      'Bus Stop', 'Business Service', 'Park', 'Fish Market', 'Plaza',
      'Historic Site', 'Pub', 'Caucasian Restaurant', 'Breakfast Spot',
      'Pizza Place', 'Fish & Chips Shop', 'Wine Shop', 'Grocery Store',
      'Supermarket', 'Hotel', 'Indian Restaurant', 'Bar', 'Coffee Shop',
      'Metro Station', 'Fast Food Restaurant', 'Outdoor Sculpture',
      'Stadium', 'Chocolate Shop', 'Multiplex', 'Restaurant',
      'Sporting Goods Shop', 'Clothing Store', 'Food Court',
      'Music Venue', 'Bistro', 'Roof Deck', 'Latin American Restaurant',
      'Outlet Mall', 'Sandwich Place', 'Asian Restaurant',
      'Gym / Fitness Center', 'Ice Cream Shop', 'Hotel Bar',
      'Burger Joint', 'American Restaurant', 'Brazilian Restaurant',
      'Bubble Tea Shop', 'Sports Bar', 'Outlet Store',
      'Electronics Store', 'Pedestrian Plaza', 'Diner', 'Juice Bar',
      'Cosmetics Shop', 'Seafood Restaurant', 'Gastropub', 'Museum',
      'Turkish Restaurant', 'Gift Shop', 'Dessert Shop',
      'English Restaurant', 'Train Station', 'Sushi Restaurant',
      'Bookstore', 'Modern European Restaurant', 'Malay Restaurant',
      'Speakeasy', 'Garden', 'Film Studio', 'Used Bookstore',
      'Deli / Bodega', 'Recording Studio', 'Art Gallery',
      'Greek Restaurant', 'Health Food Store', 'Portuguese Restaurant',
      'Szechuan Restaurant', 'Dog Run', 'Steakhouse', 'Cocktail Bar',
      'Vietnamese Restaurant', 'Lounge', 'Butcher', 'Yoga Studio',
      'New American Restaurant', 'French Restaurant', 'History Museum',
      'Falafel Restaurant', 'Wine Bar', 'Mexican Restaurant',
      'Event Space', 'Concert Hall', 'Theater', 'Performing Arts Venue',
      'Scandinavian Restaurant', 'Udon Restaurant', 'Boutique',
      'Boxing Gym', 'Botanical Garden', 'Burrito Place',
      'Indie Movie Theater', 'Cycle Studio', 'Shopping Mall',
      'Thai Restaurant', 'Building', 'Salon / Barbershop',
      'Dim Sum Restaurant', 'Gaming Cafe', 'Spanish Restaurant',
      'Nightclub', 'Caribbean Restaurant', 'Korean Restaurant',
      'Brewery', 'Donut Shop', 'Insurance Office', 'Salad Place',
      'Organic Grocery', 'Ramen Restaurant',
      'Southern / Soul Food Restaurant', 'Polish Restaurant',
      'Persian Restaurant', 'Video Game Store', 'Department Store',
      'Hookah Bar', 'Pharmacy', 'Discount Store', 'Optical Shop',
      'Stationery Store', "Women's Store", 'Irish Pub', 'Shoe Store',
      'Outdoor Supply Store', 'Beer Bar', 'Shipping Store',
      'Movie Theater', 'Paper / Office Supplies Store',
```

We are only interested in Restaurants, so we will extract any venue with the word 'Restaurant'.

```
[ 'Mediterranean Restaurant',
  'Italian Restaurant',
  'Caucasian Restaurant',
  'Indian Restaurant',
  'Fast Food Restaurant',
  'Restaurant',
  'Latin American Restaurant',
  'Asian Restaurant',
  'American Restaurant',
  'Brazilian Restaurant',
  'Seafood Restaurant',
  'Turkish Restaurant',
  'English Restaurant',
  'Sushi Restaurant',
  'Modern European Restaurant',
  'Malay Restaurant',
  'Greek Restaurant',
  'Portuguese Restaurant',
  'Szechuan Restaurant',
  'Vietnamese Restaurant',
  'New American Restaurant',
  'French Restaurant',
  'Falafel Restaurant',
  'Mexican Restaurant',
  'Scandinavian Restaurant',
  'Udon Restaurant',
  'Thai Restaurant',
  'Dim Sum Restaurant',
  'Spanish Restaurant',
  'Caribbean Restaurant',
  'Korean Restaurant',
  'Ramen Restaurant',
  'Southern / Soul Food Restaurant',
```

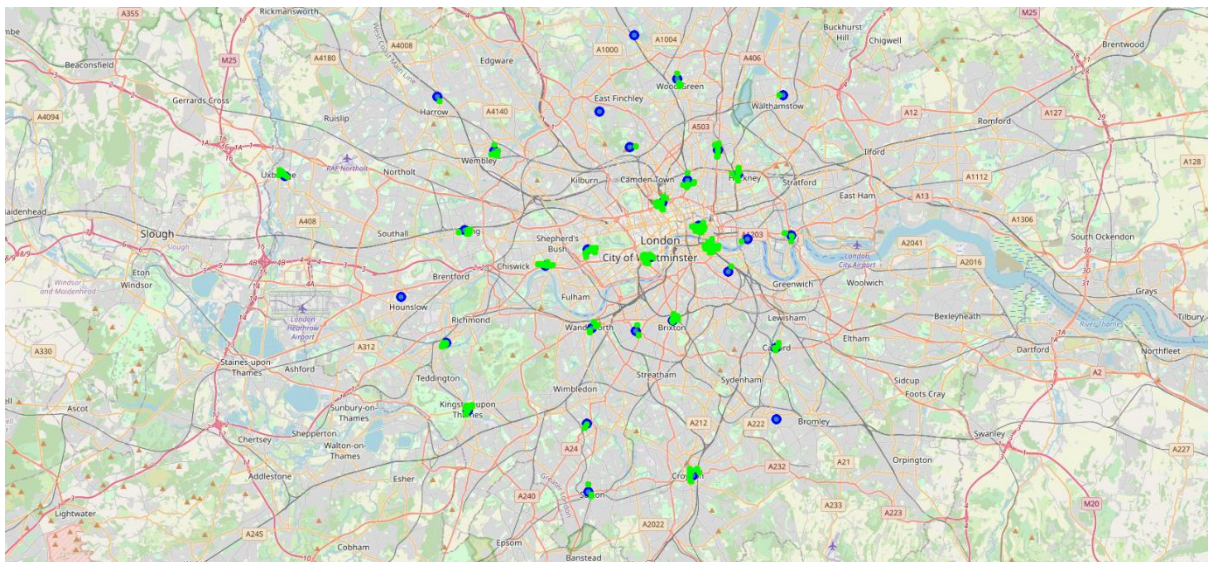
Then we put that back into our dataframe.

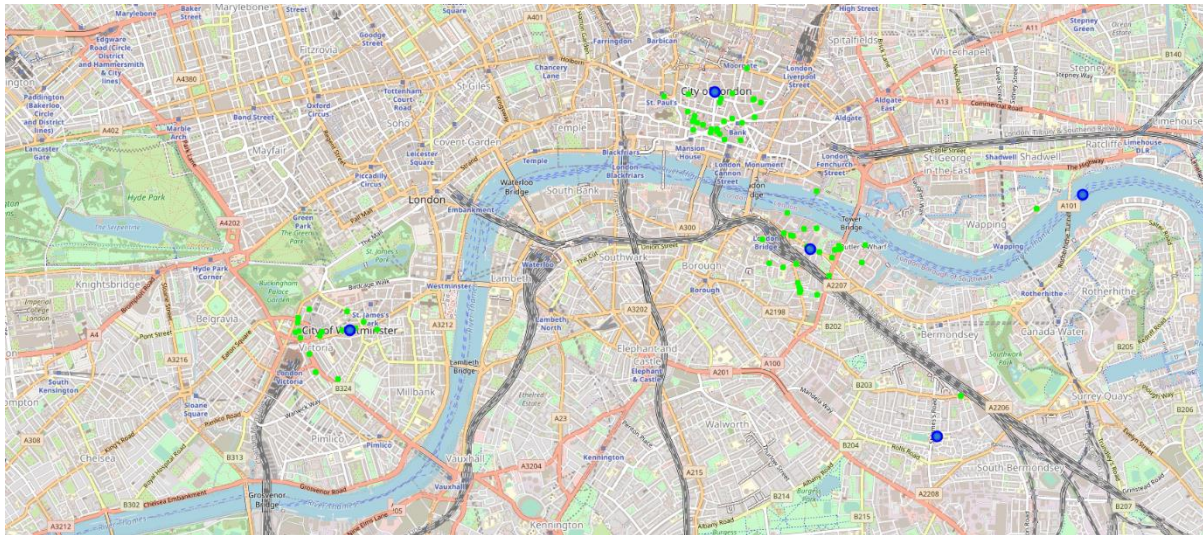
	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Barking and Dagenham	51.5607	-0.1557	Bistro Laz	51.561333	-0.150528	Mediterranean Restaurant
1	Barking and Dagenham	51.5607	-0.1557	Al Parco	51.561356	-0.150680	Italian Restaurant
2	Barking and Dagenham	51.5607	-0.1557	Carob Tree	51.561377	-0.150580	Mediterranean Restaurant
3	Bexley	51.4549	-0.1505	The Georgian	51.452242	-0.147677	Caucasian Restaurant
4	Bexley	51.4549	-0.1505	La Baita	51.458228	-0.148801	Italian Restaurant
5	Bexley	51.4549	-0.1505	Chatkhara	51.452637	-0.147465	Indian Restaurant
6	Bexley	51.4549	-0.1505	Chicken Cottage	51.452127	-0.147835	Fast Food Restaurant
7	Brent	51.5588	-0.2817	Bobby Moore Club	51.557140	-0.279610	Restaurant
8	Brent	51.5588	-0.2817	Wembley Tandoori Restaurant	51.562063	-0.281572	Indian Restaurant
9	Brent	51.5588	-0.2817	Las Iguanas	51.556443	-0.283967	Latin American Restaurant
10	Brent	51.5588	-0.2817	wagamama	51.556667	-0.283224	Asian Restaurant
11	Brent	51.5588	-0.2817	Frankie & Benny's	51.556744	-0.283913	American Restaurant
12	Brent	51.5588	-0.2817	TGI Fridays	51.557116	-0.281986	American Restaurant
13	Brent	51.5588	-0.2817	Cabana Brazilian Barbecue	51.556884	-0.282407	Brazilian Restaurant
14	Brent	51.5588	-0.2817	Prezzo	51.556717	-0.282753	Italian Restaurant
15	Brent	51.5588	-0.2817	Zizzi	51.556748	-0.283492	Italian Restaurant
16	Brent	51.5588	-0.2817	Bobby Moore Spice	51.558537	-0.278152	Indian Restaurant
17	Brent	51.5588	-0.2817	Nets Seafood Restaurant	51.556071	-0.279021	Seafood Restaurant
18	Camden	51.5290	-0.1255	Pitted Olive	51.526369	-0.125623	Turkish Restaurant
19	Camden	51.5290	-0.1255	The Gilbert Scott	51.529442	-0.126071	English Restaurant
20	Camden	51.5290	-0.1255	Mai Sushi	51.529314	-0.129701	Sushi Restaurant

Let us see how many of each restaurant there are, shall we!?

Venue Category	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude
Italian Restaurant	51	51	51	51	51	51
Indian Restaurant	29	29	29	29	29	29
Restaurant	18	18	18	18	18	18
Fast Food Restaurant	15	15	15	15	15	15
Sushi Restaurant	14	14	14	14	14	14
Portuguese Restaurant	14	14	14	14	14	14
French Restaurant	13	13	13	13	13	13
Vietnamese Restaurant	12	12	12	12	12	12
Thai Restaurant	12	12	12	12	12	12
Asian Restaurant	10	10	10	10	10	10
Japanese Restaurant	9	9	9	9	9	9
English Restaurant	9	9	9	9	9	9
Turkish Restaurant	8	8	8	8	8	8
Mediterranean Restaurant	8	8	8	8	8	8
Mexican Restaurant	7	7	7	7	7	7
Caribbean Restaurant	7	7	7	7	7	7
Modern European Restaurant	7	7	7	7	7	7
Tapas Restaurant	6	6	6	6	6	6
Vegetarian / Vegan Restaurant	6	6	6	6	6	6
Latin American Restaurant	5	5	5	5	5	5
Chinese Restaurant	5	5	5	5	5	5
Brazilian Restaurant	4	4	4	4	4	4

3.1.4 Plot the restaurant venues to our map of London





3.1.5 Explore frequency of venue category per Borough

Let us now analyse each borough with dummy coding to discover each Borough's top restaurant category frequency, so that we can then begin clustering.

Neighborhood	African Restaurant	American Restaurant	Argentinian Restaurant	Asian Restaurant	Australian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Cantonese Restaurant	Caribbean Restaurant	Caucasian Restaurant	Chinese Restaurant	Dim Sum Restaurant	Dumpling Restaurant	English Restaurant	Falafel Restaurant	Fast Food Restaurant	French Restaurant	German Restaurant	Greek Restaurant	Indian Chinese Restaurant	Indian Restaurant	Italian Restaurant
0 Barking and Dagenham	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1 Barking and Dagenham	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
2 Barking and Dagenham	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3 Bexley	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
4 Bexley	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
5 Bexley	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
6 Bexley	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
7 Brent	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8 Brent	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
9 Brent	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10 Brent	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11 Brent	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12 Brent	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13 Brent	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14 Brent	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
15 Brent	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
16 Brent	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
17 Brent	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18 Camden	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19 Camden	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
20 Camden	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Next, we group rows by neighbourhood and by taking the mean of the frequency of occurrence of each category.

Neighborhood	African Restaurant	American Restaurant	Argentinian Restaurant	Asian Restaurant	Australian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Castonese Restaurant	Caribbean Restaurant	Caucasian Restaurant	Chinese Restaurant	Dim Sum Restaurant	Dumpling Restaurant	English Restaurant	Falafel Restaurant	Fast Food Restaurant	French Restaurant	German Restaurant	Greek Restaurant	Indian Chinese Restaurant	Indian Restaurant	Italian Restaurant
0 Barking and Dagenham	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	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	0.000000	0.333333
1 Bexley	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.25	0.000000	0.000000	0.000000	0.000000	0.000000	0.250000	0.000000	0.000000	0.000000	0.000000	0.250000	0.250000
2 Brent	0.000000	0.181818	0.000000	0.090909	0.000000	0.090909	0.000000	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	0.181818	0.181818
3 Camden	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.083333	0.000000	0.000000	0.000000	0.000000	0.083333	0.000000	0.083333	0.166667
4 City of London	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.000000	0.037037	0.000000	0.037037	0.037037	0.000000	0.111111	0.000000	0.000000	0.000000	0.074074	0.111111
5 Croydon	0.000000	0.000000	0.000000	0.071429	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.071429	0.000000	0.071429
6 Ealing	0.000000	0.000000	0.000000	0.050000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.050000	0.000000	0.050000	0.000000	0.050000	0.000000	0.050000	0.000000	0.100000
7 Enfield	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.142857	0.000000	0.000000	0.000000	0.000000	0.142857	0.142857
8 Greenwich	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	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	0.000000	0.000000
9 Hackney	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	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	0.000000	0.125000
10 Hammersmith and Fulham	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.076923	0.000000	0.000000	0.038462	0.000000	0.038462	0.000000	0.000000	0.000000	0.038462	0.000000	0.153846
11 Haringey	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.285714	0.000000	0.000000	0.000000	0.000000	0.000000	0.142857
12 Harrow	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	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	0.000000	0.000000
13 Hillingdon	0.000000	0.000000	0.000000	0.090909	0.000000	0.090909	0.000000	0.000000	0.000000	0.00	0.090909	0.000000	0.000000	0.000000	0.000000	0.181818	0.000000	0.000000	0.000000	0.000000	0.090909	0.363636
14 Islington	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	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	0.142857	0.000000
15 Kensington and Chelsea	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.083333	0.000000	0.000000	0.083333	0.000000	0.000000	0.000000	0.083333	0.166667
16 Kingston upon Thames	0.000000	0.000000	0.000000	0.043478	0.000000	0.000000	0.043478	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.043478	0.043478	0.086957	0.043478	0.043478	0.000000	0.000000	0.000000	0.173913
17 Lambeth	0.029412	0.000000	0.000000	0.000000	0.000000	0.058824	0.000000	0.029412	0.147059	0.00	0.000000	0.029412	0.029412	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.058824	0.058824
18 Lewisham	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.166667	0.000000	0.000000	0.166667	0.000000	0.000000	0.333333
19 Merton	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.00	0.000000	0.000000	0.000000	0.000000	0.000000	0.333333	0.000000	0.000000	0.000000	0.000000	0.333333	0.333333
20 Newham	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	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	0.000000	1.000000
21 Redbridge	0.000000	0.000000	0.000000	0.000000	0.071429	0.000000	0.000000	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	0.071429	0.142857

Let's print each neighbourhood along with the top 10 most common venues with the frequency distribution.

```

----Barking and Dagenham----
      venue  freq
0  Mediterranean Restaurant 0.67
1      Italian Restaurant 0.33
2  Scandinavian Restaurant 0.00
3      Mexican Restaurant 0.00
4  Middle Eastern Restaurant 0.00
5  Modern European Restaurant 0.00
6  New American Restaurant 0.00
7  Okonomiyaki Restaurant 0.00
8      Persian Restaurant 0.00
9      Polish Restaurant 0.00
    
```

```

----Bexley----
      venue  freq
0  Fast Food Restaurant 0.25
1      Italian Restaurant 0.25
2      Indian Restaurant 0.25
3  Caucasian Restaurant 0.25
4      African Restaurant 0.00
5  Modern European Restaurant 0.00
6  New American Restaurant 0.00
7  Okonomiyaki Restaurant 0.00
8      Persian Restaurant 0.00
9      Polish Restaurant 0.00
    
```

```

----Brent----
      venue  freq
0  American Restaurant 0.18
1      Italian Restaurant 0.18
2      Indian Restaurant 0.18
3      Asian Restaurant 0.09
4  Brazilian Restaurant 0.09
5  Latin American Restaurant 0.09
6      Restaurant 0.09
7  Seafood Restaurant 0.09
8      African Restaurant 0.00
9  Scandinavian Restaurant 0.00
    
```


Then we put that into a pandas dataframe to show the ten most common venues per Borough.

	Neighborhood	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	Mediterranean Restaurant	Italian Restaurant	Vietnamese Restaurant	English Restaurant	Kebab Restaurant	Japanese Restaurant	Indian Restaurant	Indian Chinese Restaurant	Greek Restaurant	German Restaurant
1	Bexley	Italian Restaurant	Indian Restaurant	Caucasian Restaurant	Fast Food Restaurant	Vietnamese Restaurant	English Restaurant	Kebab Restaurant	Japanese Restaurant	Indian Chinese Restaurant	Greek Restaurant
2	Brent	American Restaurant	Italian Restaurant	Indian Restaurant	Restaurant	Brazilian Restaurant	Seafood Restaurant	Latin American Restaurant	Asian Restaurant	Cajun / Creole Restaurant	Fast Food Restaurant
3	Camden	Italian Restaurant	Szechuan Restaurant	English Restaurant	Restaurant	Modern European Restaurant	Sushi Restaurant	Malay Restaurant	Greek Restaurant	Turkish Restaurant	Indian Restaurant
4	City of London	Italian Restaurant	French Restaurant	Restaurant	Seafood Restaurant	Modern European Restaurant	Indian Restaurant	Vietnamese Restaurant	Sushi Restaurant	English Restaurant	Latin American Restaurant
5	Croydon	Portuguese Restaurant	Indian Restaurant	Spanish Restaurant	Italian Restaurant	Restaurant	Caribbean Restaurant	Mediterranean Restaurant	Malay Restaurant	Sushi Restaurant	Greek Restaurant
6	Ealing	Vietnamese Restaurant	Italian Restaurant	Polish Restaurant	Caribbean Restaurant	Greek Restaurant	French Restaurant	Fast Food Restaurant	English Restaurant	New American Restaurant	Persian Restaurant
7	Enfield	Turkish Restaurant	Restaurant	Italian Restaurant	Indian Restaurant	Fast Food Restaurant	Portuguese Restaurant	Dumpling Restaurant	Japanese Restaurant	Indian Chinese Restaurant	Greek Restaurant
8	Greenwich	Vietnamese Restaurant	Latin American Restaurant	Kebab Restaurant	Japanese Restaurant	Italian Restaurant	Indian Restaurant	Indian Chinese Restaurant	Greek Restaurant	German Restaurant	French Restaurant
9	Hackney	Vietnamese Restaurant	Vegetarian / Vegan Restaurant	Modern European Restaurant	Italian Restaurant	Ramen Restaurant	Korean Restaurant	Japanese Restaurant	Indian Restaurant	Indian Chinese Restaurant	Greek Restaurant
10	Hammersmith and Fulham	Indian Restaurant	Italian Restaurant	Vietnamese Restaurant	Chinese Restaurant	Japanese Restaurant	Tapas Restaurant	Portuguese Restaurant	Korean Restaurant	Indian Chinese Restaurant	Mediterranean Restaurant

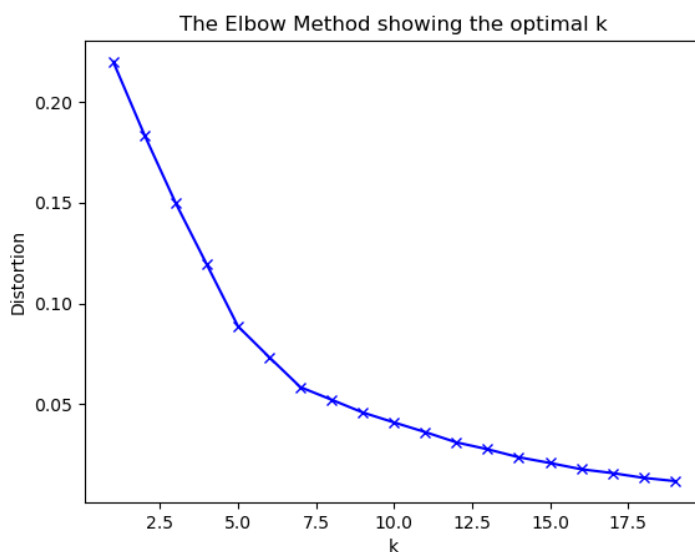
3.2 Clustering

Run k-means to cluster the neighborhood into k clusters. K-Means algorithm is one of the most common cluster method of unsupervised learning.

3.2.1 Cluster Optimisation

Analyze the K-Means with the elbow method to ensured optimum k of the K-Means.

There are different spatial distance function to experiment with^{xvi}. I have opted to for sqeuclidean (k=7, Dist=0.05) as it gives a high enough k and low enough distortion. Since London had 32 Boroughs my intuition tells me that a k higher than 5 would be a better fit.



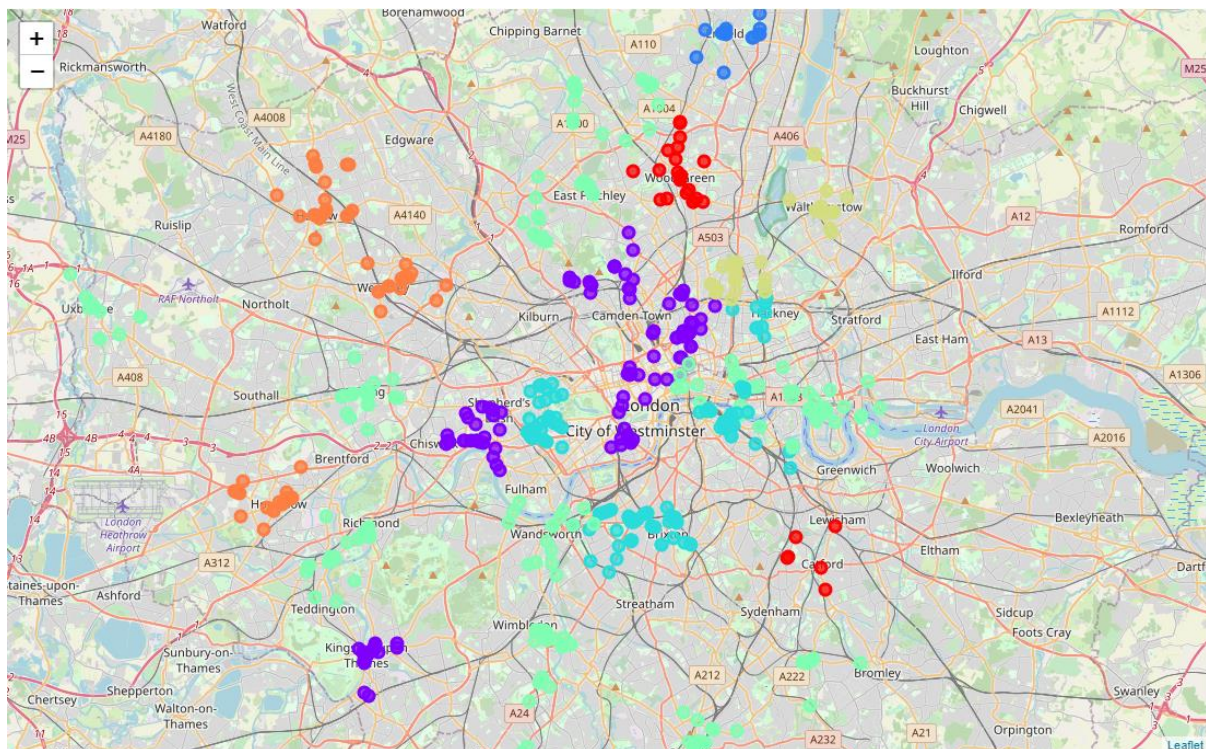
^{xvi} <https://docs.scipy.org/doc/scipy-0.16.0/reference/generated/scipy.spatial.distance.pdist.html>

We create a new dataframe that includes the cluster as well as the top 10 venues for each neighborhood.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category	Cluster Labels	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	51.5607	-0.1557	Bistro Laz	51.561333	-0.150528	Mediterranean Restaurant	0	Mediterranean Restaurant	Italian Restaurant	Vietnamese Restaurant	English Restaurant	Kebab Restaurant	Japanese Restaurant	Indian Restaurant	Indian Chinese Restaurant	Greek Restaurant	German Restaurant
1	Barking and Dagenham	51.5607	-0.1557	Al Parco	51.561356	-0.150680	Italian Restaurant	0	Mediterranean Restaurant	Italian Restaurant	Vietnamese Restaurant	English Restaurant	Kebab Restaurant	Japanese Restaurant	Indian Restaurant	Indian Chinese Restaurant	Greek Restaurant	German Restaurant
2	Barking and Dagenham	51.5607	-0.1557	Carob Tree	51.561377	-0.150580	Mediterranean Restaurant	0	Mediterranean Restaurant	Italian Restaurant	Vietnamese Restaurant	English Restaurant	Kebab Restaurant	Japanese Restaurant	Indian Restaurant	Indian Chinese Restaurant	Greek Restaurant	German Restaurant
3	Besley	51.4549	-0.1505	The Georgian	51.452242	-0.147677	Caucasian Restaurant	0	Italian Restaurant	Indian Restaurant	Caucasian Restaurant	Fast Food Restaurant	Vietnamese Restaurant	English Restaurant	Kebab Restaurant	Japanese Restaurant	Indian Chinese Restaurant	Greek Restaurant
4	Besley	51.4549	-0.1505	La Baita	51.458228	-0.148801	Italian Restaurant	0	Italian Restaurant	Indian Restaurant	Caucasian Restaurant	Fast Food Restaurant	Vietnamese Restaurant	English Restaurant	Kebab Restaurant	Japanese Restaurant	Indian Chinese Restaurant	Greek Restaurant

3.2.2 Cluster Mapping

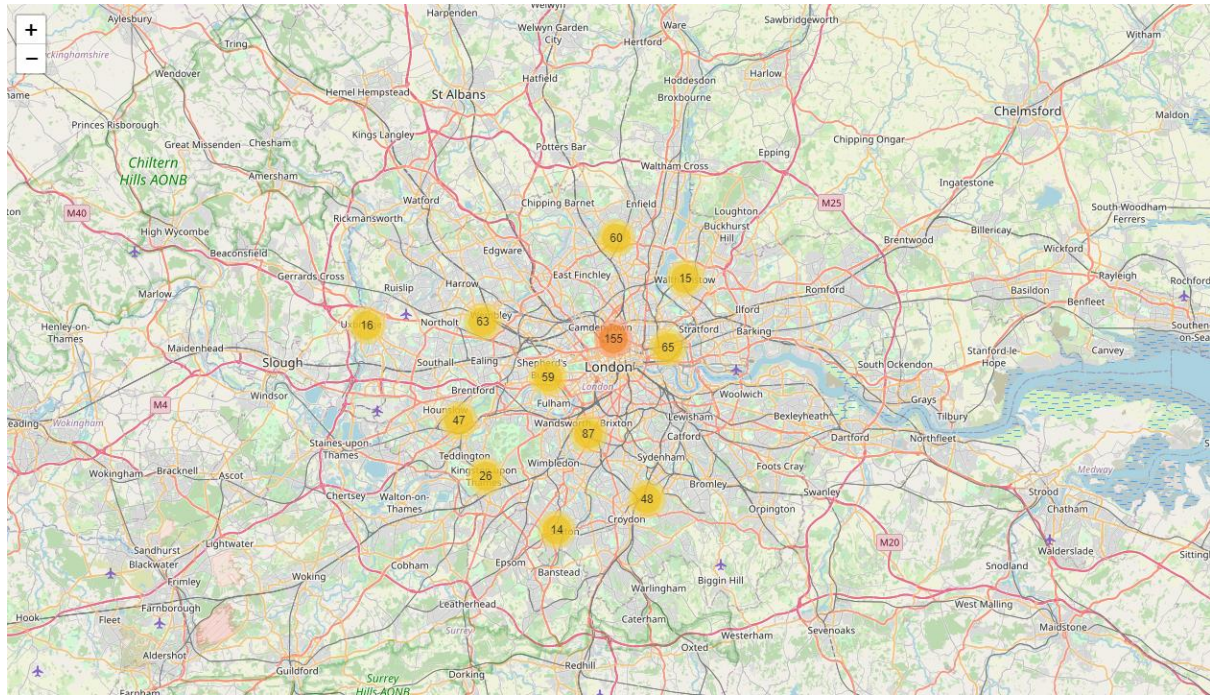
Then we visualise the dataframe.



3.2.3 Marker Clustering

The clusters look a bit congested and it's not the best data viz. To remedy this, we're going to group the markers into different clusters. Each cluster is then represented by the number of restaurants in each neighbourhood. These clusters can be thought of as pockets of London which you can then analyse separately.

To implement this, we start off by instantiating a MarkerCluster object and adding all the data points in the dataframe to this object.

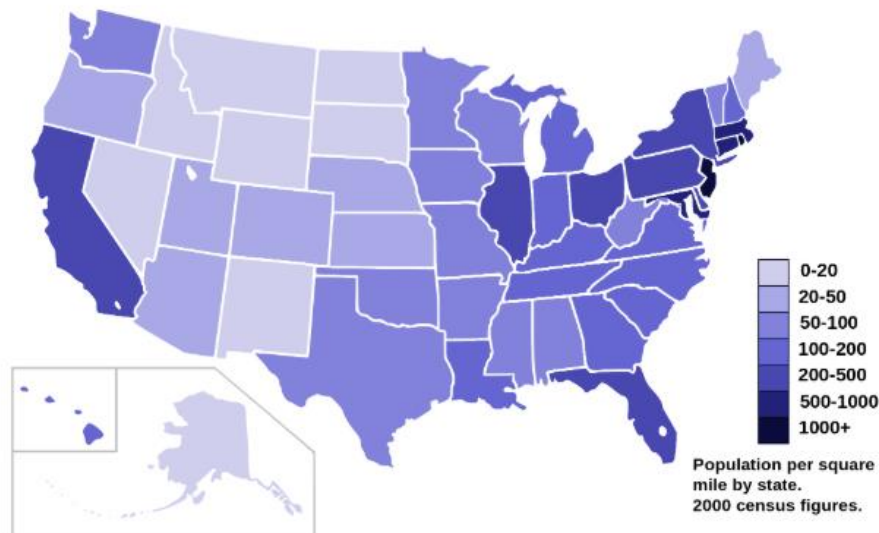


4. Results

To help us explore the results we will employ choropleth maps.

4.1 Choropleth Maps

A `Choropleth` map is a thematic map in which areas are shaded or patterned in proportion to the measurement of the statistical variable being displayed on the map, such as population density or per-capita income. The choropleth map provides an easy way to visualize how a measurement varies across a geographic area or it shows the level of variability within a region. Below is a `Choropleth` map of the US depicting the population by square mile per state.



Here I will conduct exploratory analysis even more creatively by running choropleth maps over the London Boroughs by different top restaurant categories to see which areas are "hottest" for what cuisine.

Then, I will try and do the same for opening hours (IF I have enough calls left to make to the API since these are Premium Calls). This should hopefully indicate which restaurants are still operating during the coronavirus lockdown - An indication of what cuisines are in demand enough for the restaurants to maintain operations. Correlation does NOT equal causation.. but it's reasonable to assume that if businesses are open for trade that there is demand there, its economics 101 my good sir!

In order to create a `Choropleth` map, we need a GeoJSON file that defines the areas/boundaries of the state, county, or country that we are interested in. In our case, since we are endeavouring to create a map of good old London, we want a GeoJSON that defines the boundaries of all London Boroughs. Basically, this is a set of coordinates that outline the boundary of each borough to form a polygon.

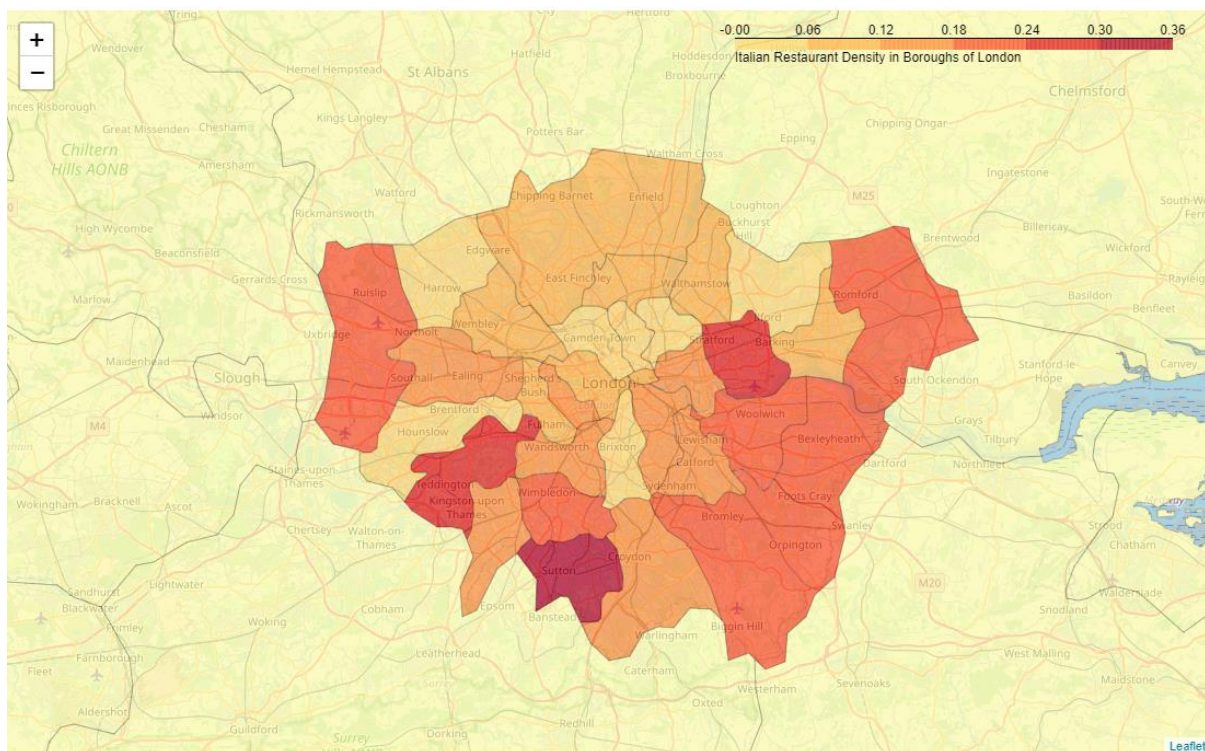
We will need the normalised dataframe from earlier which has the mean frequency of each restaurant per Neighborhood, as this is what we will use to generate choropleth maps for different cuisines.

To create a `Choropleth` map, we will use the choropleth method with the following main parameters:

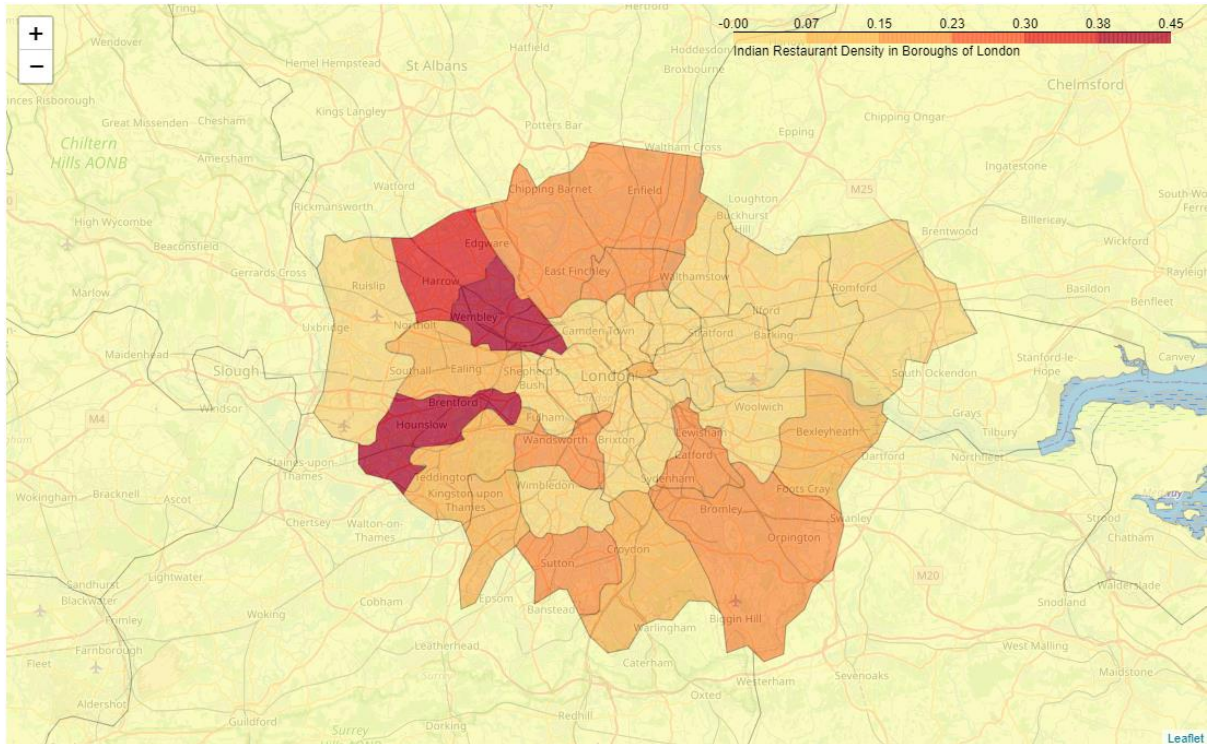
1. `geo_data`, which is the GeoJSON file.
2. `data`, which is the dataframe containing the data.
3. `columns`, which represents the columns in the dataframe that will be used to create the `Choropleth` map.
4. `key_on`, which is the key or variable in the GeoJSON file that contains the name of the variable of interest. To determine that, you will need to open the GeoJSON file using any text editor and note the name of the key or variable that contains the Borough name, since the Boroughs are our variable of interest. In this case, `**name**` is the key in the GeoJSON file that contains the name of the countries. Note that this key is case sensitive, so you need to pass exactly as it exists in the GeoJSON file.

Now we will create a choropleth map of London Boroughs for the Top 5 Restaurants by Venue Category. Ignoring 'Restaurant' as it is too generic.

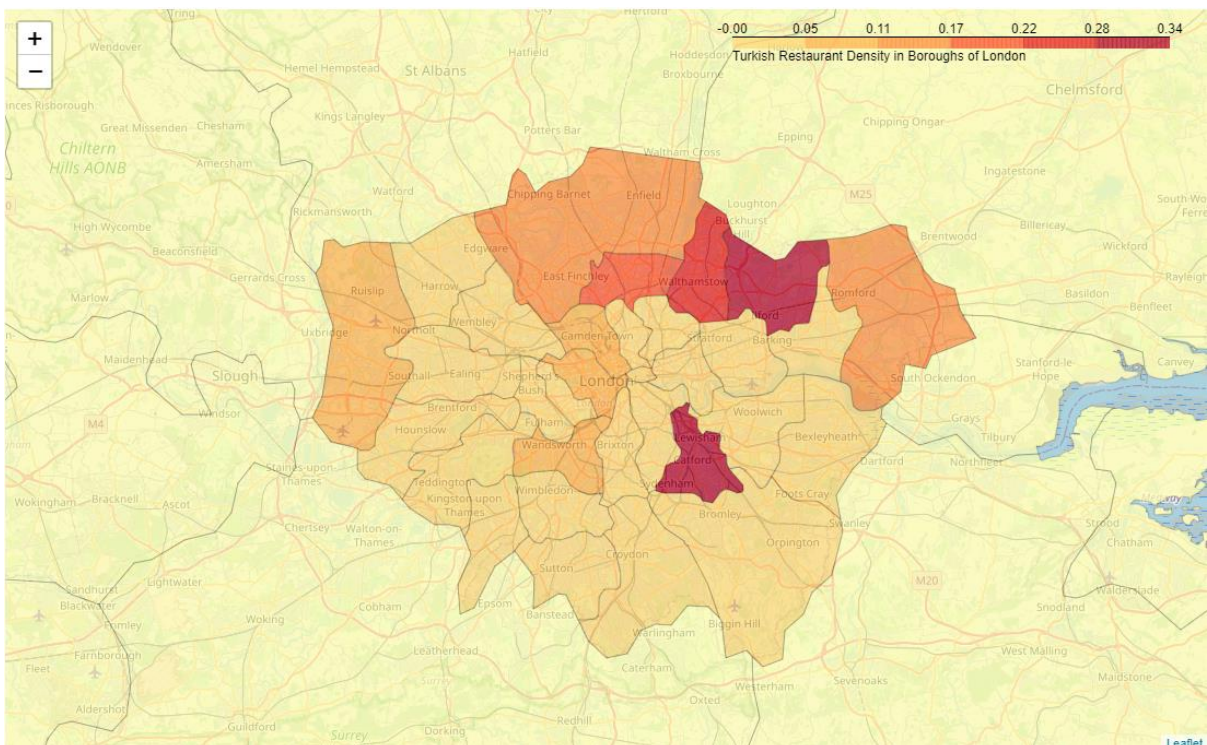
Neighborhood	
Venue Category	
Italian Restaurant	84
Indian Restaurant	68
Restaurant	35
Turkish Restaurant	35
Fast Food Restaurant	31
Thai Restaurant	31



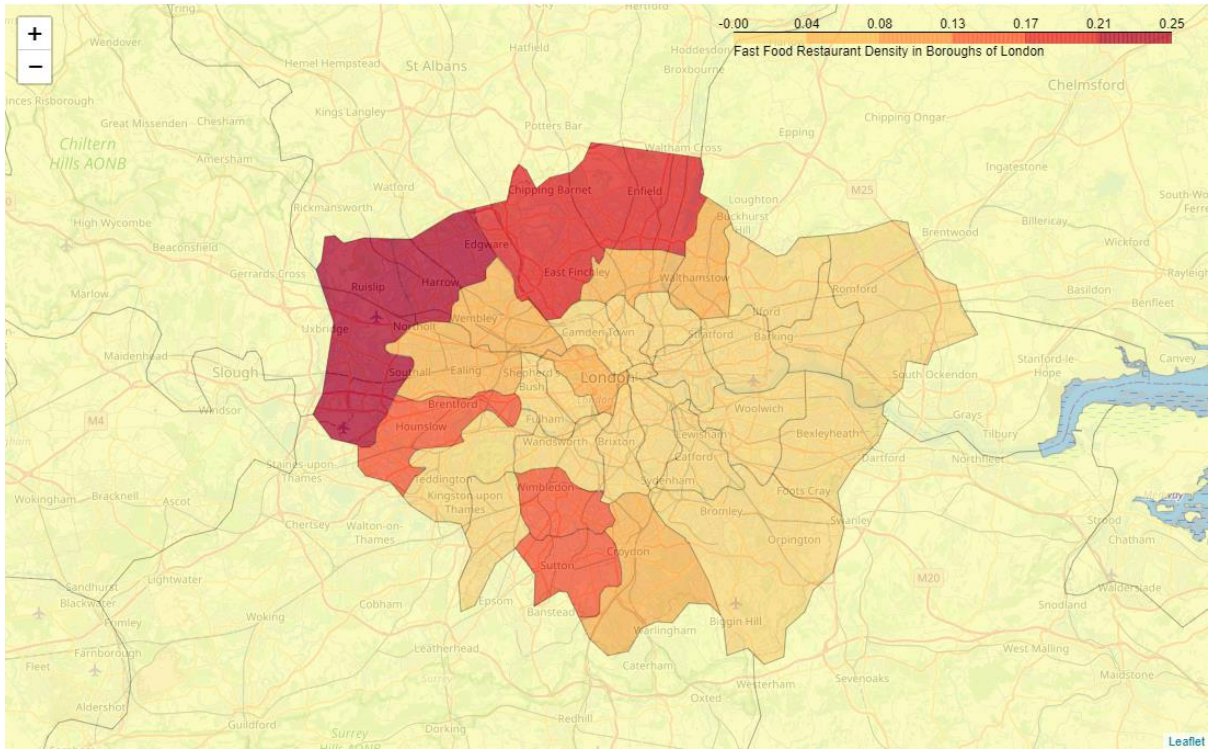
Hotspots for Italian Restaurants in London are Newham, Sutton and Richmond.



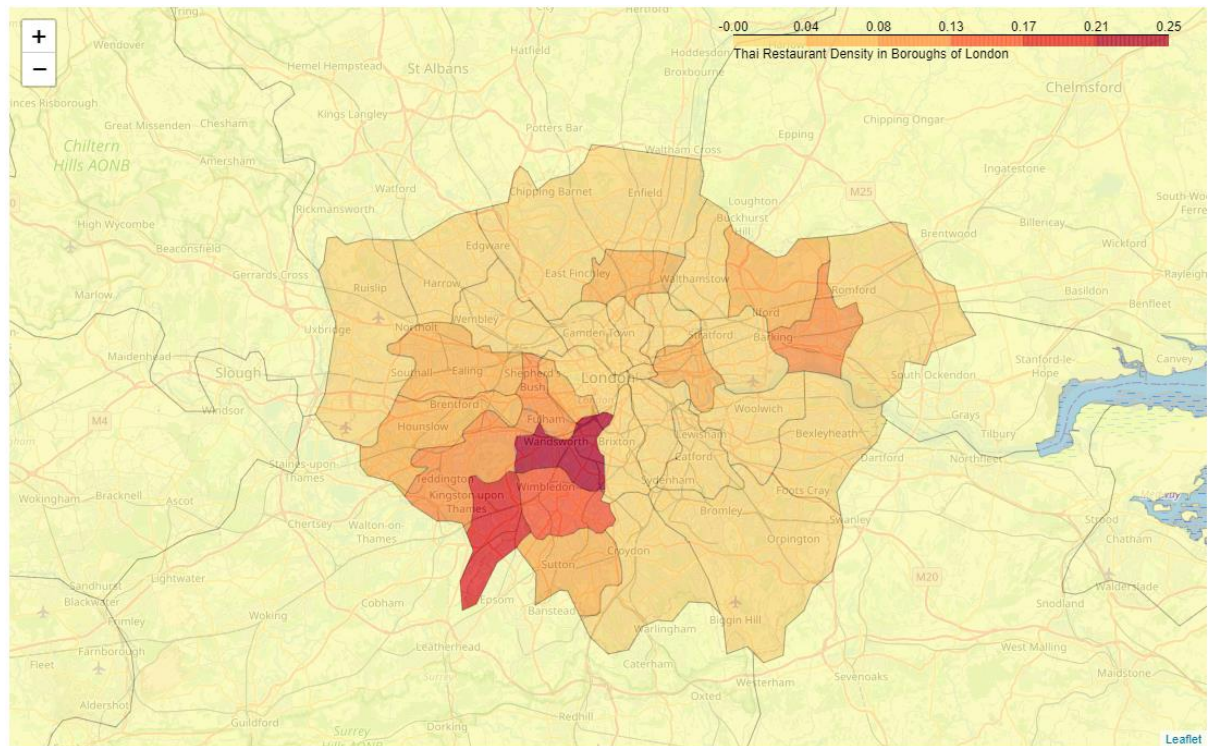
The hotspots for Indian restaurants in London are Brent and Hounslow, followed by Harrow.



The hotspot for a Turkish place in London is Lewisham and Redbridge. Waltham Forest would be a good choice too.



Fast Food joints are popular in North West London.



Thai Restaurants are popular in South West London.

5. Discussion

London has an area of 607 square miles. In comparison to NYC that has an overall area of 468.4 square miles, but, of which, just 302 is land. So, NYC has approx. half the land size as London, however they have almost identical populations; London has an estimated population in 2017 of 8.825 million; NYC has an estimated population in 2017 of 8.623 million. London is more irregular and spread out, whereas NYC has more modern design in clinical blocks, with less variation (and character/charm) and has less districts - 5 vs 32 - since it is built more vertically than London. This is predominantly due to history. The first recorded population count for New York City was 7,681 in 1698. The city grew at a moderate rate in the 18th century, but exploded in the 19th century, more than doubling in the final ten-year period. The town of 80,000 in 1800 became a city of 3.4 million by the end of the 1800s. However, London's story dates back slightly earlier: https://en.wikipedia.org/wiki/Timeline_of_London.

The restaurant clusters around London not surprisingly show that the largest cluster centres around central London (City of London, Westminster, Kensington and Chelsea) using kmeans clustering with k value of 7 using the sqeuclidean spatial distance function. I am still learning about clustering and spatial distance functions, so there will be room for improvement here. Further analysis could be done to cross reference restaurant clusters and density with ethnic density and clustering for each borough to explore the relationship between cultural preferences.

Issues I ran into: The geojson also does not split Barnet and Enfield. The FS API hours feature is a Premium call, so it is out of scope to generate choropleth maps of restaurants trading under lockdown.

Important note: This analysis has been conducted on an extremely limited dataset due to the rate limits imposed by the FS API. To present proper findings I would upgrade my account and include every restaurant in London and perform sentiment analysis of the user Tips.

6. Conclusion

Judging by the results the top most popular restaurants that I would recommend, along with the hotspot for that type of food would be; an Italian in Sutton, an Indian in Brent or Hounslow, a Turkish place in Lewisham or Redbridge, a Fast Food joint in Hillingdon/Harrow, or a Thai restaurant in Wandsworth.

The next step would be to use the estate agent API to make calls for commercial properties that are on the market in those areas for the client, but this is out of scope.