Introduction: Business Problem

In this project we will try to find an optimal location for a restaurant. Specifically, this report will be targeted to stakeholders interested in opening an **Chinese restaurant** in **Lodon**, UK.

Since there are lots of restaurants in Lodon we will try to detect **locations that are not already crowded with restaurants**. We are also particularly interested in **areas with no Chinese restaurants in vicinity**. We would also prefer locations **as close to city center as possible**, assuming that first two conditions are met.

We will use our data science powers to generate a few most promissing neighborhoods based on this criteria. Advantages of each area will then be clearly expressed so that best possible final location can be chosen by stakeholders.

Data

Based on definition of our problem, factors that will influence our decission are:

- number of existing restaurants in the neighborhood (any type of restaurant)
- number of and distance to Chinese restaurants in the neighborhood, if any
- distance of neighborhood from city center

We decided to use regularly spaced grid of locations, centered around city center, to define our neighborhoods.

Following data sources will be needed to extract/generate the required information:

- centers of candidate areas will be generated algorithmically and approximate addresses of centers of those areas will be obtained using **Google Maps API reverse geocoding**
- number of restaurants and their type and location in every neighborhood will be obtained using **Foursquare API**
- coordinate of Lodon center will be obtained using **Google Maps API geocoding** of well known London location (Lambeth North tube station)

Neighborhood Candidates

Let's create latitude & longitude coordinates for centroids of our candidate neighborhoods. We will create a grid of cells covering our area of interest which is aprox. 12x12 killometers centered around Lodon city center.

Let's first find the latitude & longitude of Lodon city center, using specific, well known address and Google Maps geocoding API.

Coordinate of Lambeth North tube station, London, UK: [51.4989177, -0.1121086]

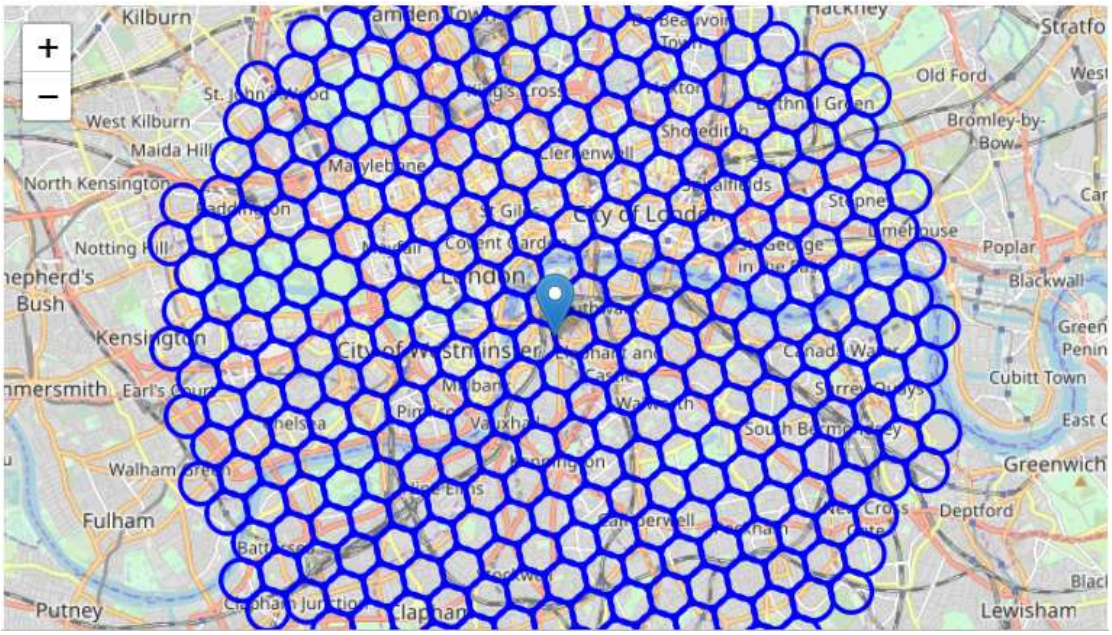
Now let's create a grid of area candidates, equaly spaced, centered around city center and within ~6km from Lambeth North tube station. Our neighborhoods will be defined as circular areas with a radius of 300 meters, so our neighborhood centers will be 600 meters apart.

To accurately calculate distances we need to create our grid of locations in Cartesian 2D coordinate system which allows us to calculate distances in meters (not in latitude/longitude degrees). Then we'll project those coordinates back to latitude/longitude degrees to be shown on Folium map. So let's create functions to convert between WGS84 spherical coordinate system (latitude/longitude degrees) and UTM Cartesian coordinate system (X/Y coordinates in meters).

Let's create a **hexagonal grid of cells**: we offset every other row, and adjust vertical row spacing so that **every cell center is equally distant from all it's neighbors**.

364 candidate neighborhood centers generated.

Let's visualize the data we have so far: city center location and candidate neighborhood centers:



OK, we now have the coordinates of centers of neighborhoods/areas to be evaluated, equally spaced (distance from every point to it's neighbors is exactly the same) and within ~6km from Lambeth North tube station.

Let's now use Google Maps API to get approximate addresses of those locations.

Reverse geocoding check

Address of [51.4989177, -0.1121086] is: Lambeth North Station, 110 Westminster Bridge Rd, South Bank, London SE1 7XG, UK

Foursquare

Now that we have our location candidates, let's use Foursquare API to get info on restaurants in each neighborhood.

We're interested in venues in 'food' category, but only those that are proper restaurants - coffe shops, pizza places, bakeries etc. are not direct competitors

so we don't care about those. So we will include in out list only venues that have 'restaurant' in category name, and we'll make sure to detect and include all the subcategories of specific 'Chinese restaurant' category, as we need info on Chinese restaurants in the neighborhood.

Total number of restaurants: 2909

Total number of Chinese restaurants: 163

Percentage of Chinese restaurants: 5.60%

Average number of restaurants in neighborhood: 7.271978021978022

List of all restaurants

('4c2e38bf7d85a593eef553f3', 'Luanda Grill', 51.44742303467162, -0.12437808709929998, '256a Brixton Hill, Brixton, Greater London, SW2 1HF, United Kingdom', 304, False, -548184.3357480518, 5808903.871061945)

('4dcd81fcd22deadedd3acafc', 'Wing Fu', 51.445970117008564, -0.12436866760253906, '282 Brixton Hill (at Morrish Rd), Brixton, Greater London, SW2 1HT, United Kingdom', 279, True, -548217.5553274208, 5808743.506958647)

('4c8769706e65199c44f75c53', 'Ten Thanks', 51.4486955597945, -0.12050062020187428, '31 Upper Tulse Hill (Ostade Rd & Wimbart Rd), Brixton, Greater London, SW2 2SD, United Kingdom', 327, True, -547887.484229397, 5808987.73223665)

('4e287183a809ec0663ee4892', 'Umana Yana', 51.45043303563944, -0.10060129055622724, '294 Croxted Road, Herne Hill, Greater London, SE24 9DA, United Kingdom', 327, False, -546475.8092566801, 5808889.763127419)

('4d84fe1081fdb1f7bc4108c0', 'Rocca', 51.44923236432431, -0.08515756266426021, '75-79 Dulwich Village, Dulwich, Greater London, SE21 7BJ, United Kingdom', 181, False, -545439.5368921933, 5808532.879800605)

('4e66ccafc65be127eea697cb', 'Surma Curry House', 51.450528, -0.079358, 'Dulwich, Greater London, S E22, United Kingdom', 118, False, -545009.776879589, 5808591.534941786)

('4cc1e61f38aaa093dfd40262', 'Bombay Bicycle Club', 51.449664, -0.074881, 'Lordship lane, Camberwell, Greater London, United Kingdom', 332, False, -544721.346755112, 5808431.246874542)

('4be31054b02ec9b6d4774ec0', 'Curry Cabin', 51.451117, -0.070042, 'Lordship lane, East Dulwich, Greater London, United Kingdom', 165, False, -544354.1427618747, 5808521.257353644)

('4fcb80d8e4b00865ccd6d4c9', 'Oriental Star', 51.45192840689823, -0.06723108981546579, 'United Kingdom', 226, True, -544141.6054765151, 5808569.9569546105)

('506db71ee4b00eb58f69b364', 'Royal British Legion Dulwich', 51.454596, -0.069029, 'United Kingdom', 240, False, -544203.5395660708, 5808890.242570764)

...

Total: 2909

List of Chinese restaurants

('4dcd81fcd22deadedd3acafc', 'Wing Fu', 51.445970117008564, -0.12436866760253906, '282 Brixton Hill (at Morrish Rd), Brixton, Greater London, SW2 1HT, United Kingdom', 279, True, -548217.5553274208, 5808743.506958647)

('4c8769706e65199c44f75c53', 'Ten Thanks', 51.4486955597945, -0.12050062020187428, '31 Upper Tulse Hill (Ostade Rd & Wimbart Rd), Brixton, Greater London, SW2 2SD, United Kingdom', 327, True, -547887.484229397, 5808987.73223665)

('4fcb80d8e4b00865ccd6d4c9', 'Oriental Star', 51.45192840689823, -0.06723108981546579, 'United Kingdom', 226, True, -544141.6054765151, 5808569.9569546105)

('4ca66047931bb60cc0bc92e2', 'Mr Liu', 51.45538945768149, -0.07672434147241174, '58 Lordship Ln, East Dulwich, Greater London, S E22, United Kingdom', 139, True, -544715.329469735, 5809089.440604274)

('4ccb2f17c0378cfabb087d48', 'Silver Fish Bar', 51.466922221286296, -0.052050286383005434, '48 Gibbon Rd, Camberwell, Greater London, SE1 5 2, United Kingdom', 256, True, -542747.6757855524, 5810003.445119728)

('4fa44009e4b0cfe54c0e2cfb', 'Courtesan Dim Sum', 51.46115994755113, -0.11122012921135563, '69-73 Atlantic Rd, Brixton, Greater London, SW9 8PU, United Kingdom', 82, True, -546957.6561985367, 5810227.206030627)

('5884c87b8d8e99297680cce0', 'Duck Duck Goose', 51.462715, -0.112497, 'Pop Brixton (49 Brixton Station Road), London, SW9 8PQ, United Kingdom', 170, True, -547009.390986866, 5810417.27708085)

('56994b03498e09226e6d2bd1', 'Mr Bao', 51.465824, -0.065894, '293 Rye Ln, London, Greater London, SE15 4UA, United Kingdom', 326, True, -543726.7918829196, 5810083.044752231)

('4e3d2b52483b04e17a91da46', 'Lovely House', 51.46789189338897, -0.07241128097607828, '119 Bellenden Road, Camberwell, Greater London, United Kingdom', 179, True, -544127.6882852758, 5810405.659747191)

('57433804498e6896a9ac6c3d', 'On Cafe', 51.461159936690485, -0.13606309890747068, '31 Clapham Park Road, SW4 7EE, United Kingdom', 217, True, -548669.0347714156, 5810589.009541841)

...

Total: 163

Restaurants around location

Restaurants around location 101: Pho Ta

Restaurants around location 102:

Restaurants around location 103: Hot Stuff, The Three Lions, Bar Estrela, Grelha D'Ouro, Caf e Portugal

Restaurants around location 104: The Village Restaurante

Restaurants around location 105: Adulis, SW9 Sushi Bar, Sun Kong, Spring Way, Oval Tandoo ri

Restaurants around location 106: Zest of India

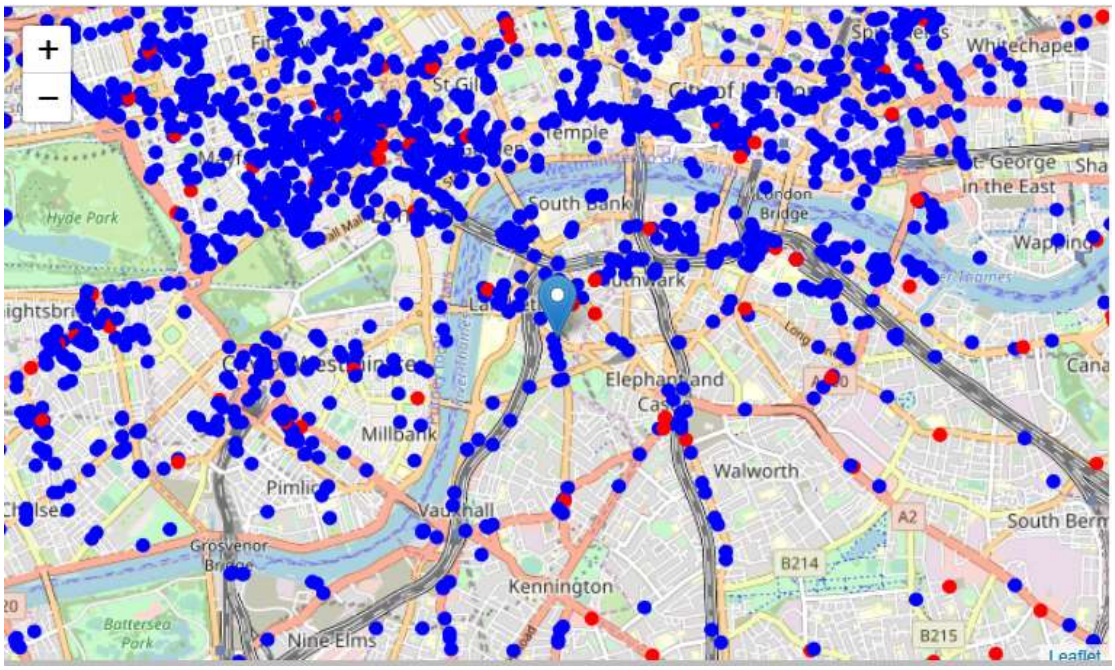
Restaurants around location 107: Bayroot, La Luna

Restaurants around location 108:

Restaurants around location 109:

Restaurants around location 110:

Let's now see all the collected restaurants in our area of interest on map, and let's also show Chinese restaurants in different color.



Looking good. So now we have all the restaurants in area within few kilometers from Lambeth North tube station, and we know which ones are Chinese restaurants! We also know which restaurants exactly are in vicinity of every neighborhood candidate center.

This concludes the data gathering phase - we're now ready to use this data for analysis to produce the report on optimal locations for a new Chinese restaurant!

Methodology

In this project we will direct our efforts on detecting areas of Lodon that have low restaurant density, particularly those with low number of Chinese restaurants. We will limit our analysis to area ~6km around city center.

In first step we have collected the required **data: location and type (category) of every restaurant within 6km from Lodon center** (Lambeth North tube station). We have also **identified Chinese restaurants**(according to Foursquare categorization).

Second step in our analysis will be calculation and exploration of '**restaurant density**' across different areas of Lodon - we will use **heatmaps** to identify a few promising areas close to center with low number of restaurants in general (*and* no Chinese restaurants in vicinity) and focus our attention on those areas.

In third and final step we will focus on most promising areas and within those create **clusters of locations that meet some basic requirements** established in discussion with stakeholders: we will take into consideration locations with **no more than two restaurants in radius of 250 meters**, and we want locations **without Chinese restaurants in radius of 400 meters**. We will present map of all such locations but also create clusters (using **k-means clustering**) of those locations to identify general zones / neighborhoods / addresses which should be a starting point for final 'street level' exploration and search for optimal venue location by stakeholders.

Analysis

Let's perform some basic explanatory data analysis and derive some additional info from our raw data. First let's count the **number of restaurants in every area candidate**:

	Address	Latitude	Longitude	X	Y	Distance from center	Restaurants in area
0	64 Holmewood Gardens, London SW2 3NB	51.446001	-0.120338	-547939.055669	5.808688e+06	5992.495307	2
1	3 Papworth Way, London SW2 2NL	51.447101	-0.112003	-547339.055669	5.808688e+06	5840.376700	0
2	151B Norwood Rd, London SE24 9AF	51.448201	-0.103668	-546739.055669	5.808688e+06	5747.173218	0
3	Edward Alleyn Club, 85 Burbage Rd, Dulwich, Lo...	51.449300	-0.095332	-546139.055669	5.808688e+06	5715.767665	0
4	29 Boxall Rd, Dulwich, London SE21 7JS	51.450399	-0.086995	-545539.055669	5.808688e+06	5747.173218	1
5	68B Beauval Rd, Dulwich, London SE22 8UQ	51.451497	-0.078658	-544939.055669	5.808688e+06	5840.376700	1
6	149A Barry Rd, East Dulwich, London SE22 0JP	51.452594	-0.070321	-544339.055669	5.808688e+06	5992.495307	3
7	58 Plummer Rd, London SW4 8HH	51.448859	-0.134366	-548839.055669	5.809208e+06	5855.766389	1
8	206 Lyham Rd, London SW2 5NR	51.449961	-0.126031	-548239.055669	5.809208e+06	5604.462508	3

Average number of restaurants in every area with radius=300m: 7.271978021978022

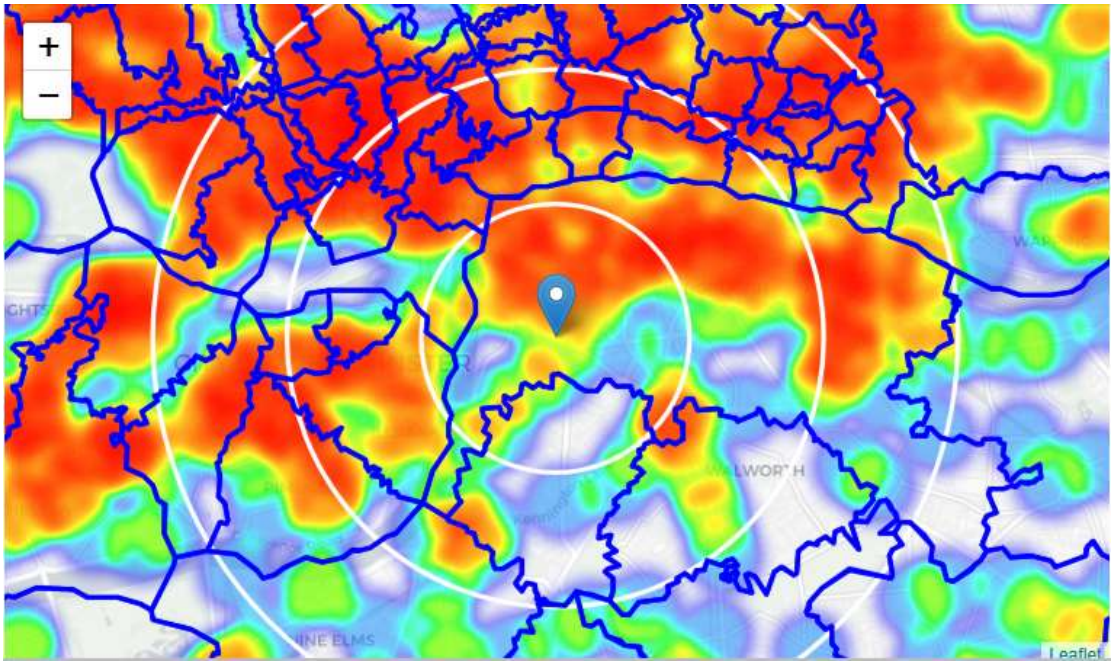
OK, now let's calculate the **distance to nearest Chinese restaurant from every area candidate center** (not only those within 300m - we want distance to closest one, regardless of how distant it is).

	Address	Latitude	Longitude	X	Y	Distance from center	Restaurants in area	Distance to Chinese restaurant
0	64 Holmewood Gardens, London SW2 3NB	51.446001	-0.120338	-547939.055669	5.808688e+06	5992.495307	2	283.940215
1	3 Papworth Way, London SW2 2NL	51.447101	-0.112003	-547339.055669	5.808688e+06	5840.376700	0	624.899629
2	151B Norwood Rd, London SE24 9AF	51.448201	-0.103668	-546739.055669	5.808688e+06	5747.173218	0	1186.850377
3	Edward Alleyn Club, 85 Burbage Rd, Dulwich, Lo...	51.449300	-0.095332	-546139.055669	5.808688e+06	5715.767665	0	1479.188454
4	29 Boxall Rd, Dulwich, London SE21 7JS	51.450399	-0.086995	-545539.055669	5.808688e+06	5747.173218	1	916.257084
5	68B Beauval Rd, Dulwich, London SE22 8UQ	51.451497	-0.078658	-544939.055669	5.808688e+06	5840.376700	1	459.407885
6	149A Barry Rd, East Dulwich,	51.452594	-0.070321	-544339.055669	5.808688e+06	5992.495307	3	230.142540

Average distance to closest Chinese restaurant from each area center: 500.6189653717583

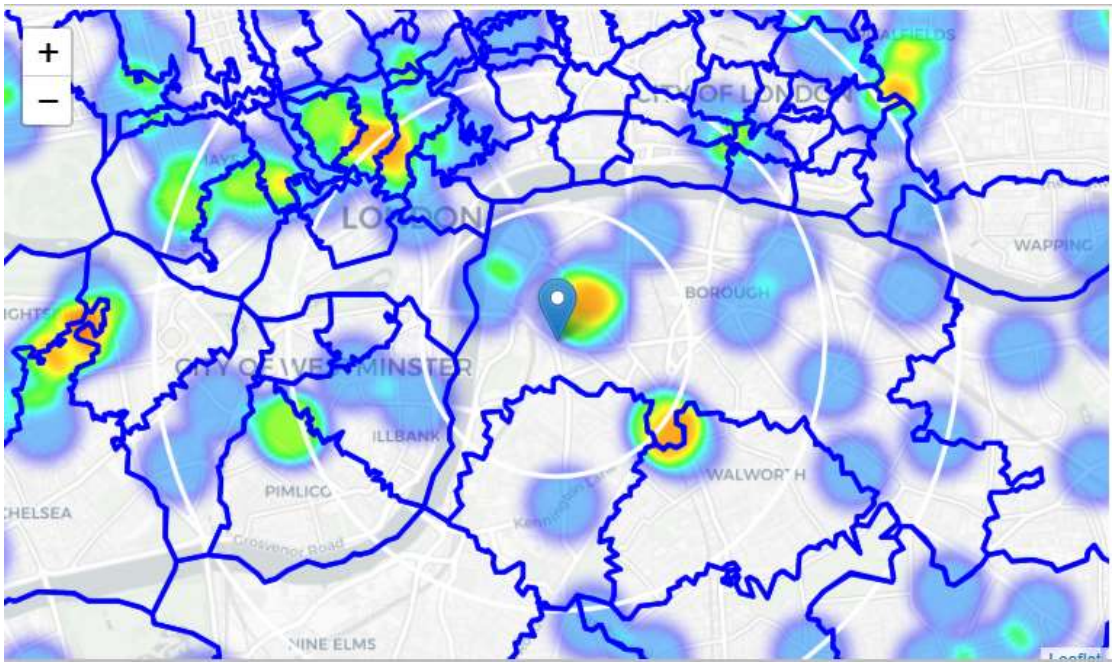
OK, so **on average Chinese restaurant can be found within ~500m** from every area center candidate. That's fairly close, so we need to filter our areas carefully!

Let's crete a map showing **heatmap / density of restaurants** and try to extract some meaningfull info from that. Also, let's show **borders of Lodon boroughs** on our map and a few circles indicating distance of 1km, 2km and 3km from Lambeth North tube station.



Looks like a few pockets of low restaurant density closest to city center can be found **south, south-east and east from Lodon center**.

Let's create another heatmap map showing **heatmap/density of Chinese restaurants** only.



This map is not so 'hot' (Chinese restaurants represent a subset of ~15% of all restaurants in Lodon) but it also indicates higher density of existing Chinese restaurants directly north and west from Sydney Tower, with closest pockets of **low Chinese restaurant density positioned east, south-east and south from city center**.

Based on this we will now focus our analysis on areas *south-west, south, south-east and east from Lodon center* - we will move the center of our area of interest and reduce it's size to have a radius of **2.5km**. This places our location candidates mostly in boroughs **Walworth and Kennington**.

Walworth and Kennington

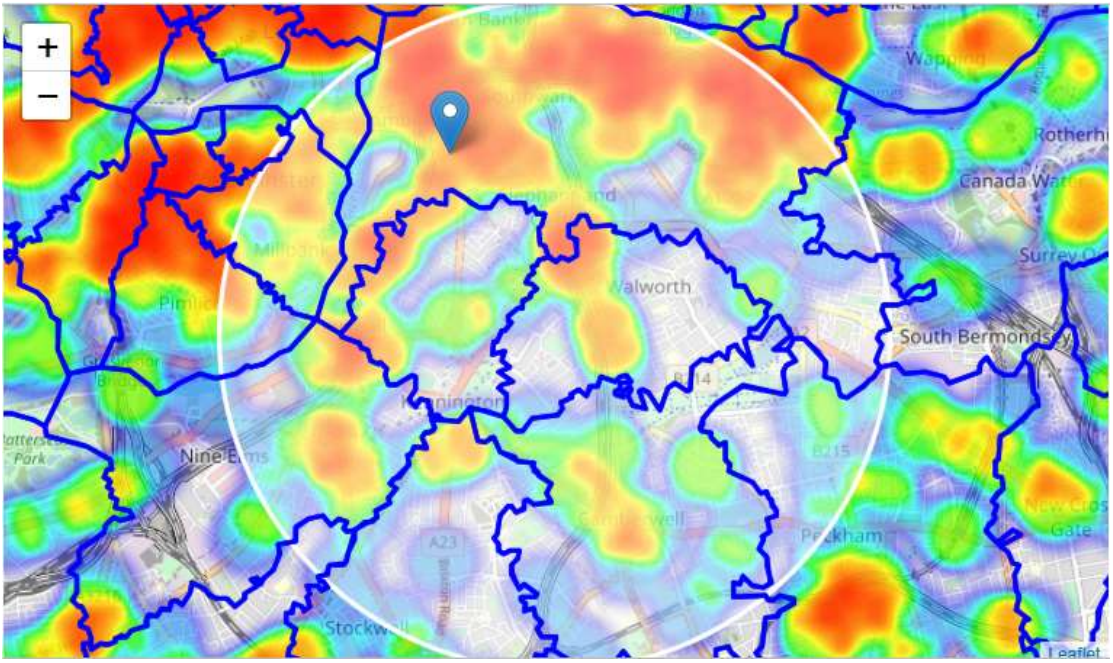
Analysis of popular travel guides and web sites often mention Walworth and Kennington as beautifull, interesting, rich with culture, 'hip' and 'cool' Lodon neighborhoods popular with tourists and loved by Lodoners.

"*Walworth has long been revered for its diverse cultural life and as a part of Lodon where alternative lifestyles have flourished. Envisioning the glamorous yet gritty nature of Lodon often conjures up scenes from this neighbourhood, where cultures, movements and artistic flare adorn the walls of building and fills the air. Brimming with nightclubs, street food, and art galleries, Walworth is the place to be for Lodon's young and trendy.*" (theculturetrip.com)

"As anyone from Walworth will tell you, this district is not just the coolest in Lodon, but the hippest location in the entire universe. Walworth has long been famed for its diverse cultural life, its experimental alternative lifestyles and the powerful spell it exercises on young people from across Germany. In 2001, Walworth and Kennington were merged to form one administrative borough. When it comes to club culture, Kennington is now out in front – with southern Kennington particularly ranked as home to the highest density of clubs in the city."

Popular with tourists, alternative and bohemian but booming and trendy, relatively close to city center and well connected, those boroughs appear to justify further analysis.

Let's define new, more narrow region of interest, which will include low-restaurant-count parts of Walworth and Kennington closest to Lambeth North tube station.



Not bad - this nicely covers all the pockets of low restaurant density in Walworth and Kennington closest to Lodon center.

Let's also create new, more dense grid of location candidates restricted to our new region of interest (let's make our location candidates 100m appart).

2261 candidate neighborhood centers generated.

OK. Now let's calculate two most important things for each location candidate: **number of restaurants in vicinity** (we'll use radius of **250 meters**) and **distance to closest Chinese restaurant**.

	Latitude	Longitude	X	Y	Restaurants nearby	Distance to Chinese restaurant
0	51.465019	-0.094109	-545689.055669	5.810404e+06	0	636.649300
1	51.465202	-0.092719	-545589.055669	5.810404e+06	0	610.555881
2	51.464763	-0.102007	-546239.055669	5.810491e+06	0	560.766961
3	51.464947	-0.100617	-546139.055669	5.810491e+06	1	549.005559
4	51.465130	-0.099227	-546039.055669	5.810491e+06	1	555.296878
5	51.465313	-0.097837	-545939.055669	5.810491e+06	1	579.052798
6	51.465496	-0.096447	-545839.055669	5.810491e+06	1	618.263424
7	51.465679	-0.095058	-545739.055669	5.810491e+06	1	576.767449
8	51.465863	-0.093668	-545639.055669	5.810491e+06	0	538.624956
9	51.466046	-0.092278	-545539.055669	5.810491e+06	0	517.274585

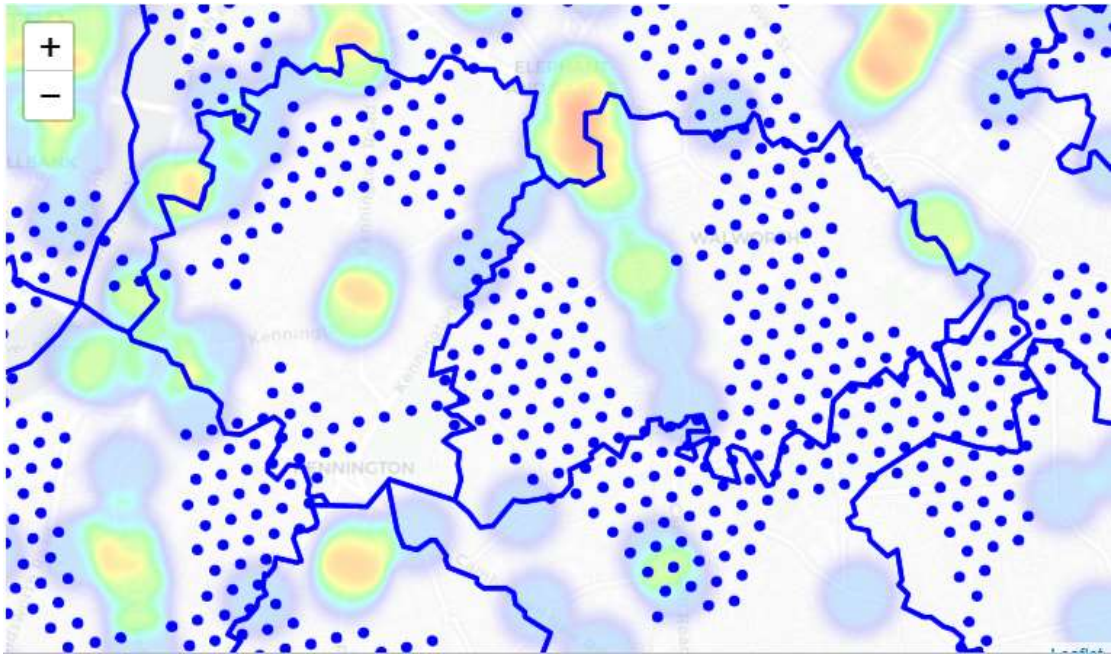
OK. Let us now **filter** those locations: we're interested only in **locations with no more than two restaurants in radius of 250 meters**, and **no Chinese restaurants in radius of 400 meters**.

Locations with no more than two restaurants nearby: 1429

Locations with no Chinese restaurants within 400m: 1152

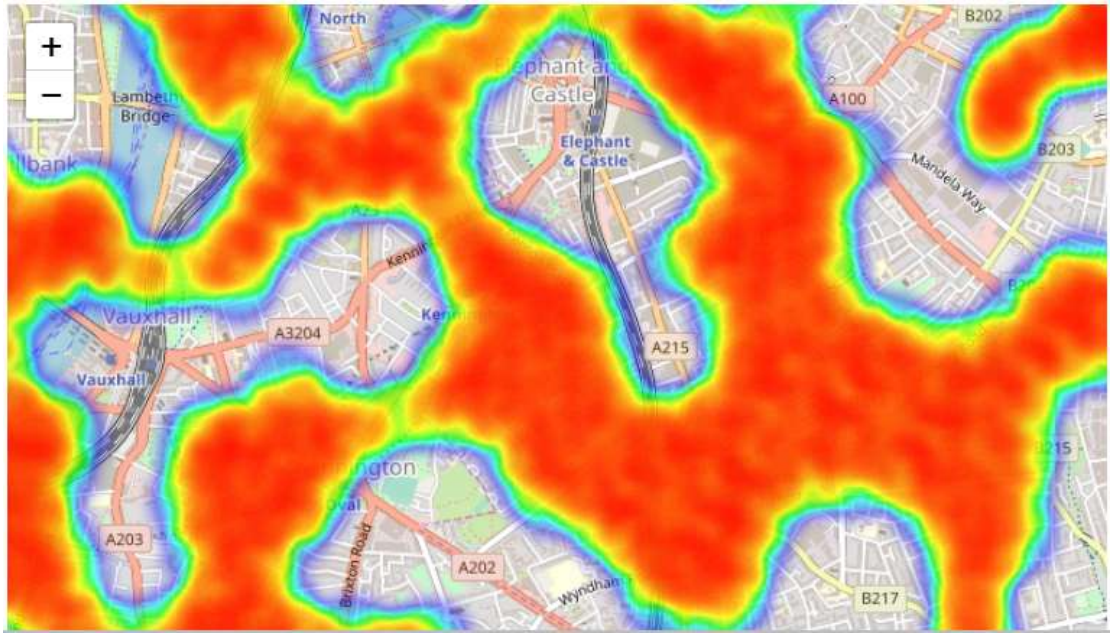
Locations with both conditions met: 852

Let's see how this looks on a map.



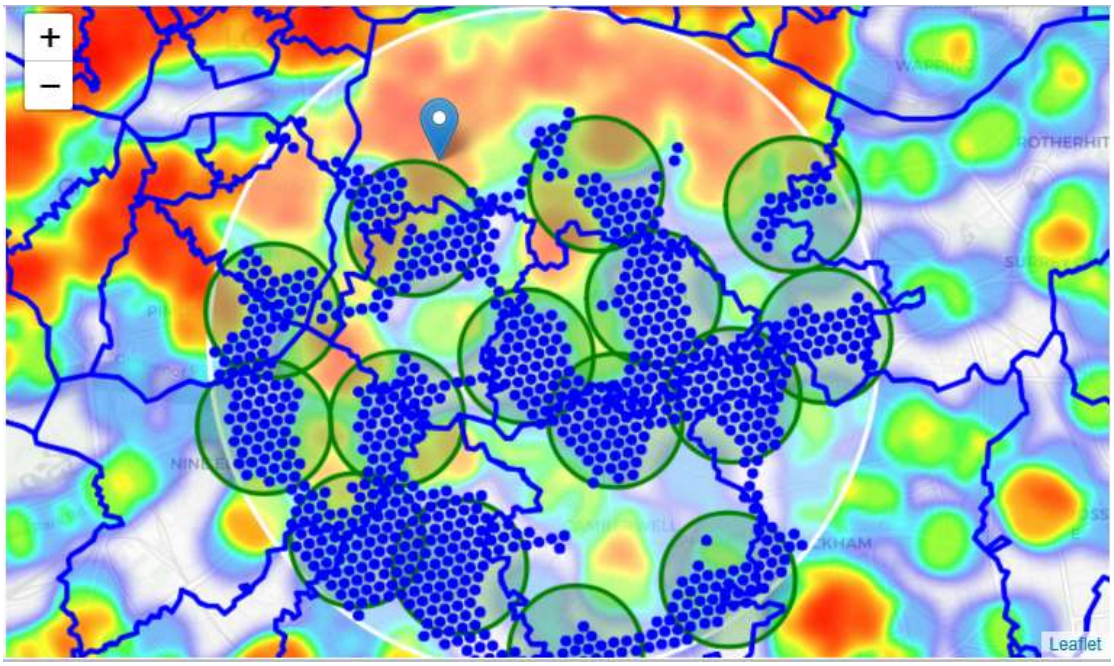
Looking good. We now have a bunch of locations fairly close to Lambeth North tube station (mostly in Walworth and Kennington), and we know that each of those locations has no more than two restaurants in radius of 250m, and no Chinese restaurant closer than 400m. Any of those locations is a potential candidate for a new Chinese restaurant, at least based on nearby competition.

Let's now show those good locations in a form of heatmap:



Looking good. What we have now is a clear indication of zones with low number of restaurants in vicinity, and *no* Chinese restaurants at all nearby.

Let us now **cluster** those locations to create **centers of zones containing good locations**. Those zones, their centers and addresses will be the final result of our analysis.



Not bad - our clusters represent groupings of most of the candidate locations and cluster centers are placed nicely in the middle of the zones 'rich' with location candidates.

Addresses of those cluster centers will be a good starting point for exploring the neighborhoods to find the best possible location based on neighborhood specifics.
Finaly, let's **reverse geocode those candidate area centers to get the addresses** which can be presented to stakeholders.

```
============================================================
====
Addresses of centers of areas recommended for further analysis
============================================================
====
63 Claylands Rd, London SW8 1PJ                    => 2.0km from Lambeth North
Tube Station
Burgess Park detras de Canchas Sintetica, 7QH, Albany Rd, London, United Kingdom =>
 2.8km from Lambeth North Tube Station
King George IV Court, 153 East St, London SE17 2SB => 1.9km from Lambeth North T
ube Station
189 Clapham Rd, London SW9 0QE                     => 2.9km from Lambeth Nort
h Tube Station
113 Camberwell Rd, London SE5 0DH                  => 2.4km from Lambeth North
 Tube Station
```

9 Ponton Rd, London SW8 5BA => 2.4km from Lambeth North Tube Station

51 Westcott Rd, London SE17 3QY => 1.6km from Lambeth North Tube Station

1 Bessborough Gardens, Westminster, London SW1V 2JQ => 1.7km from Lambeth North Tube Station

19 Spa Rd, London SE16 3SA => 2.7km from Lambeth North Tube Station

22 Northway Rd, Brixton, London SE5 9AN => 3.8km from Lambeth North Tube Station

21 Oswyth Rd, Camberwell, London SE5 8NH => 3.8km from Lambeth North Tube Station

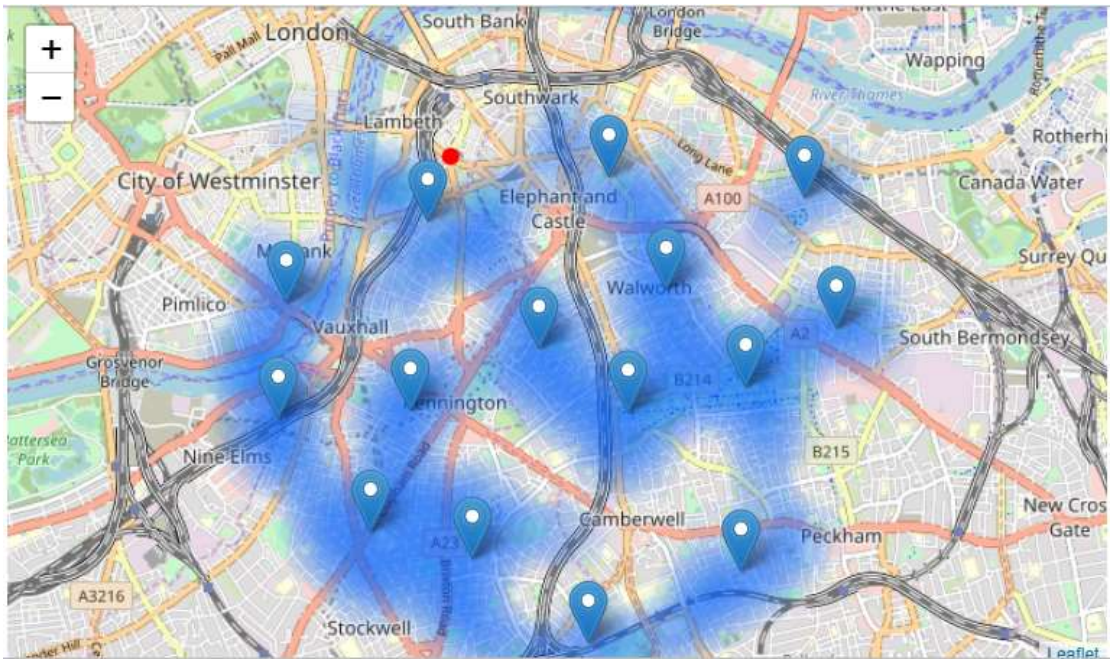
73 Lambeth Walk, Lambeth, London SE11 6DX => 0.5km from Lambeth North Tube Station

Hope Court, Fortune Pl, London SE1 5JP => 3.2km from Lambeth North Tube Station

6 Bath Terrace, London SE1 6QE => 1.2km from Lambeth North Tube Station

14 St Lawrence Way, Myatts Field South, London SW9 6NP => 3.1km from Lambeth North Tube Station

This concludes our analysis. We have created 15 addresses representing centers of zones containing locations with low number of restaurants and no Chinese restaurants nearby, all zones being fairly close to city center (all less than 4km from Lambeth North tube station, and about half of those less than 2km from Lambeth North tube station). Although zones are shown on map with a radius of ~500 meters (green circles), their shape is actually very irregular and their centers/addresses should be considered only as a starting point for exploring area neighborhoods in search for potential restaurant locations. Most of the zones are located in Walworth and Kennington boroughs, which we have identified as interesting due to being popular with tourists, fairly close to city center and well connected by public transport.



Results and Discussion

Our analysis shows that although there is a great number of restaurants in Lodon (~2000 in our initial area of interest which was 12x12km around Lambeth North tube station), there are pockets of low restaurant density fairly close to city center. Highest concentration of restaurants was detected north and west from Lambeth North tube station, so we focused our attention to areas south, south-east and east, corresponding to boroughs Walworth, Kennington and south-east corner of central Mitte borough. Another borough was identified as potentially interesting, but our attention was focused on Walworth and Kennington which offer a combination of popularity among tourists, closeness to city center, strong socio-economic dynamics *and* a number of pockets of low restaurant density.

After directing our attention to this more narrow area of interest (covering approx. 5x5km south-east from Lambeth North tube station) we first created a dense grid of location candidates (spaced 100m appart); those locations were then filtered so that those with more than two restaurants in radius of 250m and those with an Chinese restaurant closer than 400m were removed.

Those location candidates were then clustered to create zones of interest which contain greatest number of location candidates. Addresses of centers of those zones were also generated using reverse geocoding to be used as markers/starting points for more detailed local analysis based on other factors.

Result of all this is 15 zones containing largest number of potential new restaurant locations based on number of and distance to existing venues - both restaurants in general and Chinese restaurants particularly. This, of course, does not imply that those zones are actually optimal locations for a new restaurant! Purpose of this analysis was to only provide info on areas close to Lodon center but not crowded with existing restaurants (particularly Chinese) - it is entirely possible that there is a very good reason for small number of restaurants in any of those areas, reasons which would make them unsuitable for a new restaurant regardless of lack of competition in the area. Recommended zones should therefore be considered only as a starting point for more detailed analysis which could eventually result in location which has not only no nearby competition but also other factors taken into account and all other relevant conditions met.

Conclusion

Purpose of this project was to identify Lodon areas close to center with low number of restaurants (particularly Chinese restaurants) in order to aid stakeholders in narrowing down the search for optimal location for a new Chinese restaurant. By calculating restaurant density distribution from Foursquare data we have first identified general boroughs that justify further analysis (Walworth and Kennington), and then generated extensive collection of locations which satisfy some basic requirements regarding existing nearby restaurants. Clustering of those locations was then performed in order to create major zones of interest (containing greatest number of potential locations) and addresses of those zone centers were created to be used as starting points for final exploration by stakeholders.

Final decission on optimal restaurant location will be made by stakeholders based on specific characteristics of neighborhoods and locations in every

recommended zone, taking into consideration additional factors like attractiveness of each location (proximity to park or water), levels of noise / proximity to major roads, real estate availability, prices, social and economic dynamics of every neighborhood etc.