

Predicting the optimal location for a bike shop in London

Bucky BJ ZHENG March 20, 2021

1. Introduction

1.1 Background

To be fit and healthy you need to be physically active. Regular physical activity can help protect you from serious diseases such as obesity, heart disease, cancer, mental illness, diabetes and arthritis. Riding your bicycle regularly is one of the best ways to reduce your risk of health problems associated with a sedentary lifestyle.

Cycling is a healthy, low-impact exercise that can be enjoyed by people of all ages, from young children to older adults. It is also fun, cheap and good for the environment.

Riding to work or the shops is one of the most time-efficient ways to combine regular exercise with your everyday routine. An estimated one billion people ride bicycles every day – for transport, recreation and sport.

1.2 Problem

In this project we will try to find an optimal location for a bike shop. Specifically, this report will be targeted to stakeholders interested in opening an Road Bike Shop in London, UK.

Since there are lots of bike shop in London we will try to detect locations that are not already crowded with bike shop. We are also particularly interested in areas with no Road Bike Shop 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.

2. Data acquisition and cleaning

2.1 Data sources

Based on definition of our problem, factors that will influence our decision are, number of existing bike shop in the neighborhood (any type of restaurant), number of and distance to road bike shop 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 London center will be obtained using Google Maps API geocoding of well known Chelsea, London

2.2 Data collecting

We create latitude & longitude coordinates for centroids of our candidate neighborhoods. We will create a grid of cells covering our area of interest which is approx. 12x12 kilometers centered around London Chelsea center. Let's first find the latitude & longitude of London city center, using specific, well known address and Google Maps geocoding API.

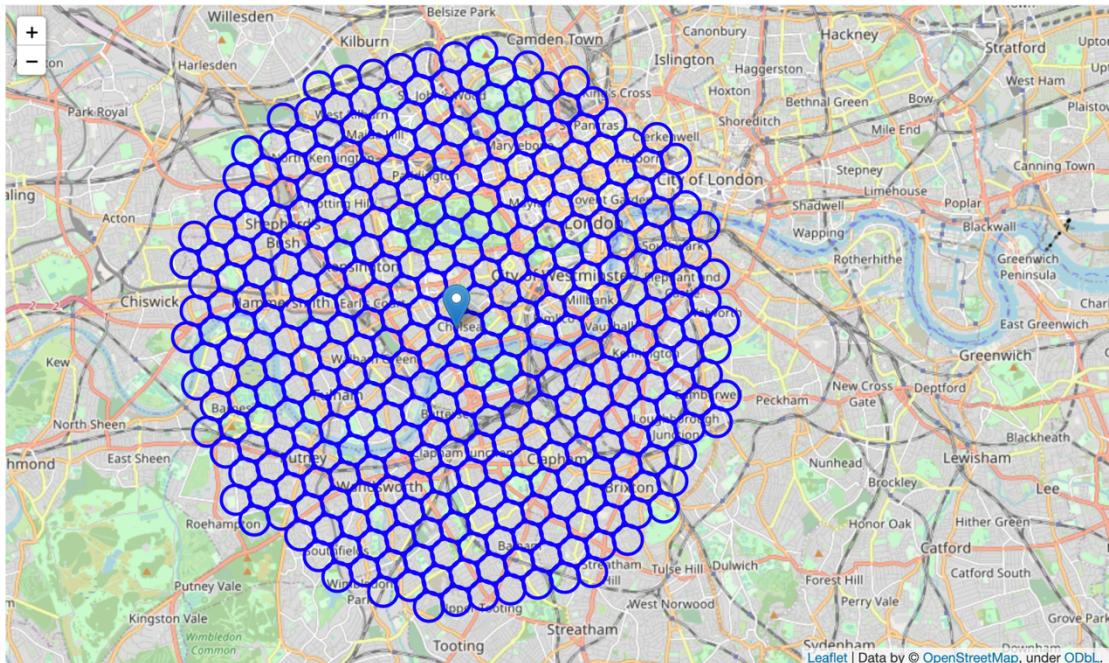
Now let's create a grid of area candidates, equally spaced, centered around city center and within ~6km from Chelsea. Our neighborhoods will be defined as circular areas with a radius of 300 meters, so our neighborhood centers will be 600 meters apart.

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).

Then, we got coordinate of Chelsea, London, UK: [51.486943, -0.170037], and also got coordinate of London centre longitude is -0.1700369999999926, latitude is 51.48694299999999. Then 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.

Then, 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 and visualize the data we have so far: city center location and candidate neighborhood centers:



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 Chelsea. Now use Google Maps API to get approximate addresses of those locations, and we got 364 related locations.

```
[ '89 Strathdon Dr, London SW17 0PR',
  '4 John Hunter Ave, London SW17 7FA',
  '36 Noyna Rd, Tooting, London SW17 7PH',
  '12 Manville Rd, Balham, London SW17 8JN',
  '15 Brierley Rd, London SW12 9LY',
  "5 Mortimer Cl, St Leonard's, London SW16 1AQ",
  '15 Killieser Ave, Streatham Hill, London SW2 4NX',
  'unit 15, 55-59 Weir Rd, London SW19 8UG',
  '49 Waldron Rd, Earlsfield, London SW18 3TA',
  'Clubhouse, Burntwood Ln, London SW17 0AT',
  '25 Wandle Rd, London SW17 7DL',
  '61 Nevis Rd, Balham, London SW17 7QL',
  '20A Oakmead Rd, Balham, London SW12 9SL',
  '28A Pentney Rd, Thornton, London SW12 0NX',
  '68 Thornton Rd, Thornton, London SW12 0LF',
  '6 Streatham Pl, Brixton, London SW2 4PY',
  '21 Athlone Rd, Tulse Hill, London SW2 2DT',
  '164 Elsenham St, Wimbledon Park, London SW18 5NR',
  '174 Astonville St, Southfields, London SW18',
```

And put the data into the Pandas Dataframe and save in local:

	Address	Latitude	Longitude	X	Y	Distance from center
0	89 Strathdon Dr, London SW17 0PR	51.434028	-0.178197	-552206.234464	5.808212e+06	5992.495307
1	4 John Hunter Ave, London SW17 7FA	51.435132	-0.169867	-551606.234464	5.808212e+06	5840.376700
2	36 Noyna Rd, Tooting, London SW17 7PH	51.436236	-0.161536	-551006.234464	5.808212e+06	5747.173218
3	12 Manville Rd, Balham, London SW17 8JN	51.437339	-0.153204	-550406.234464	5.808212e+06	5715.767665
4	15 Brierley Rd, London SW12 9LY	51.438441	-0.144873	-549806.234464	5.808212e+06	5747.173218

2.3 Using Foursquare scratch the related shop data

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 'sport shop' category. So we will include in out list only venues that have 'sport shop' in category name, and we'll make sure to detect and include all the subcategories of specific 'bike shop', 'bike studio', 'bike trail' and other related categories, as we need info on bike shop in the neighborhood. Then got the JSON return data from Foursquare:

```
{
  'meta': {'code': 200, 'requestId': '6054f170fa9dc32cbfc20e89'},
  'response': {'headerLocation': 'Current map view',
  'headerFullLocation': 'Current map view',
  'headerLocationGranularity': 'unknown',
  'query': 'outdoors recreation',
  'totalResults': 4,
  'suggestedBounds': {'ne': {'lat': 51.438527851203155,
    'lng': 0.1854019405612948},
  'sw': {'lat': 51.42952784220315, 'lng': 0.17099227764557004}},
  'groups': [{ 'type': 'Recommended Places',
    'name': 'recommended',
    'items': [ { 'reasons': { 'count': 0,
      'items': [ { 'summary': 'This spot is popular',
        'type': 'general',
        'reasonName': 'globalInteractionReason'} ]},
    'venue': { 'id': '4c56e2b4738e9c74093b6d21',
      'name': 'Dartford Heath',
      'location': { 'address': 'Dartford Heath',
        'lat': 51.43546592836818,
        'lon': 0.1854019405612948} } } ] } ] }
```

We totally got 1426 sports shops in Chelsea neighborhood, total number of Italian bike shops is 31, percentage of bike shops is 2.17% and average number of sport shop in neighborhood is 3.3653846153846154.

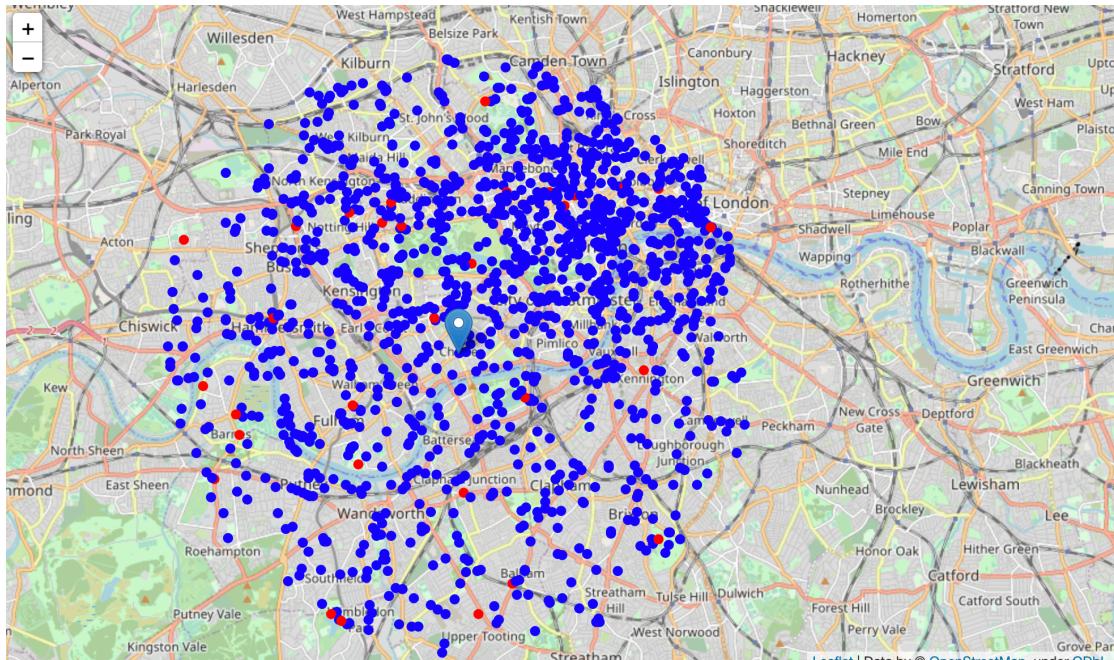
This is list of sports shops:

```
( '4c6ea34dc52437048b772beb', 'Garratt Green', 51.4356712983604, -0.18192213880395477, 'Burntwood Lane (Aboyne Road), Earlsfield, Greater London, United Kingdom', 291, False, -552424.5321573927, 5808447.821918954)
('4c44201336d6a59384046aa8', "St George's Park, Wandsworth", 51.430712765483705, -0.17484955872001862, 'Wandsworth, Greater London, United Kingdom', 436, False, -552053.028098901, 5807797.612797099)
('4e3b37a6ae6045423647f517', "Tipouf's Hang Out", 51.435513, -0.184025, 'Wandsworth, Greater London, United Kingdom', 436, False, -552573.1729298043, 5808461.122220997)
('4fae6240e4b08885ff427a84', 'Central Golf Club', 51.438032606712945, -0.17379863628958453, 'United Kingdom', 169, False, -551809.4073596043, 5808589.440154857)
('529b336011d2919b189a89f3', 'Rascals Soft Play', 51.43232276457595, -0.17409577648108998, 'London, Greater London, United Kingdom', 428, False, -551963.4183197478, 5807964.135738223)
('4bad1c5f964a520c26f3be3', 'DW Fitness First', 51.436343278037384, -0.1583555614130722, '279-291 Balham High Rd., London, Greater London, SW17 7BA, United Kingdom', 374, False, -550784.5208889802, 5808177.542346447)
('578a0b89498e4abe691f485e', 'CrossFit Tooting', 51.438032, -0.163699, '22A Langroyd Road, London, Greater London, SW17 7PL, United Kingdom', 343, True, -551113.3466081619, 5808441.802461803)
('5bccb1285780eee002b6db131', 'Revive', 51.436802, -0.15772127, '249 Balham High Road, London, Greater London, SW17 7B E, United Kingdom', 272, False, -550730.0866884566, 5808218.867071485)
('575be3f6498eb2216372a8fc', 'Upper Tooting Park Play Space', 51.439534, -0.15905, 'Upper Tooting Park, London, Greater London, United Kingdom', 196, False, -550757.8390439046, 5808539.541710816)
('4ac518f6f964a52074af20e3', 'Balham Leisure Centre', 51.4415329060553, -0.1520092379285485, 'Elmfield Road, London, Greater London, SW17 8AN, United Kingdom', 139, False, -550225.9130697416, 5808657.191613987)
...
Total: 1426
```

This is list of bike shops:

```
('578a0b89498e4abe691f485e', 'CrossFit Tooting', 51.438032, -0.163699, '22A Langroyd Road, London, Greater London, SW17 7PL, United Kingdom', 343, True, -551113.3466081619, 5808441.802461803)
('56d2db91cd10af732bbc5ee8', 'Balham Volleyball Club', 51.4437799091513, -0.15352176651714444, 'Chestnut Grove (Boundaries Road), London, Greater London, SW12 8JZ, United Kingdom', 344, True, -550277.6706453164, 5808927.057694708)
('4bd719474e32d13a79b0c380', 'Wimbledon Park Athletics Track', 51.4378869796496, -0.2082517976540998, 'Home Park Rd, Wimbledon, Greater London, SW19 7HX, United Kingdom', 247, True, -554187.296169322, 5809077.555784137)
('535e689c498efe9a86cd908', 'Wimbledon Park Sailing Base', 51.43659594898185, -0.20516725071648056, 'Home Park Rd, Wimbledon, Greater London, SW19 7HX, United Kingdom', 356, True, -554004.9727120406, 5808890.008120934)
('4faf7b01edb02f54b87f319c', 'Brockwell Park BMX Track', 51.45206101133684, -0.10901406755054754, 'United Kingdom', 216, True, -547017.5916911958, 5809191.673552735)
('5bec8d4f475abdd002cb1a02e', 'Psycole', 51.46078, -0.168061, '84 Battersea Rise, London, Greater London, SW11 1EH, United Kingdom', 334, True, -550882.1096182286, 5811014.021872601)
('4ac518f9f964a5200eb020e3', 'The Hurlingham Club', 51.4661895527859, -0.2002546068366558, 'Ranelagh Gardens, London, Greater London, SW6 3PR, United Kingdom', 328, True, -552972.9340589661, 5812081.310059022)
('4bc0b250f8219c743bfcbb110', 'Roehampton Club', 51.463378291177094, -0.2438847532438461, 'Roehampton Ln, London, Greater London, SW15 5LR, United Kingdom', 475, True, -556043.8942905348, 5812410.982079737)
('50a8dc64eb0495d7f3bf2a7', 'Running Trail in Battersea Park', 51.47885882955053, -0.14954841128432184, 'United Kingdom', 312, True, -549184.5803137047, 5812737.378856311)
('55d99ca7498edde5331862b3', 'Peter May Stand', 51.48402, -0.113664, 'United Kingdom', 134, True, -546593.3361756152, 5812783.729623766)
...
Total: 31
```

Then we use folium.Map to create a map to visualize the sports shops and bike shops, the blue points represent sports shops, the red points represent bike shops:



So now we have all the sport shops in area within few kilometers from Chelsea centre, and we know which ones are bike shops. We also know which sport shops 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 bike shops!

3. Exploratory Data Analysis

In this project we will direct our efforts on detecting areas of London that have low sport shop density, particularly those with low number of bike shops. 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 sport shops within 6km from Chelsea center. We have also identified bike shops (according to Foursquare categorization).

Second step in our analysis will be calculation and exploration of 'shop density' across different areas of London - we will use heatmaps to identify a few promising areas close to center with low number of sport shops in general (and no bike shops 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 bike shops in radius of 250 meters, and we want locations without bike shops 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.

3.1 Data Analysis

We perform some basic explanatory data analysis and derive some additional info from our raw data. First let's count the number of sport shops in every area candidate, to check which area have less related shops and better opportunities, here is the data:

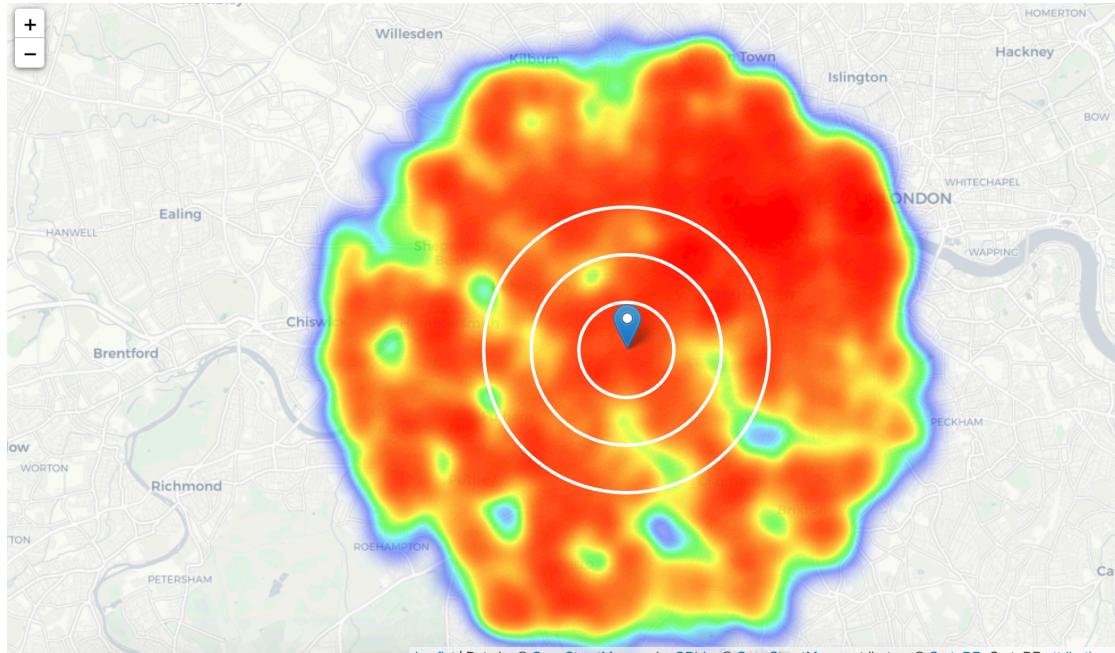
	Address	Latitude	Longitude	X	Y	Distance from center	shops in area
0	89 Strathdon Dr, London SW17 0PR	51.434028	-0.178197	-552206.234464	5.808212e+06	5992.495307	0
1	4 John Hunter Ave, London SW17 7FA	51.435132	-0.169867	-551606.234464	5.808212e+06	5840.376700	0
2	36 Noyna Rd, Tooting, London SW17 7PH	51.436236	-0.161536	-551006.234464	5.808212e+06	5747.173218	4
3	12 Manville Rd, Balham, London SW17 8JN	51.437339	-0.153204	-550406.234464	5.808212e+06	5715.767665	0
4	15 Brierley Rd, London SW12 9LY	51.438441	-0.144873	-549806.234464	5.808212e+06	5747.173218	2
5	5 Mortimer Cl, St Leonard's, London SW16 1AQ	51.439543	-0.136540	-549206.234464	5.808212e+06	5840.376700	1
6	15 Killieser Ave, Streatham Hill, London SW2 4NX	51.440645	-0.128207	-548606.234464	5.808212e+06	5992.495307	1
7	unit 15, 55-59 Weir Rd, London SW19 8UG	51.436879	-0.192223	-553106.234464	5.808732e+06	5855.766389	0
8	49 Waldron Rd, Earlsfield, London SW18 3TA	51.437985	-0.183893	-552506.234464	5.808732e+06	5604.462508	1
9	Clubhouse, Burntwood Ln, London SW17 0AT	51.439089	-0.175562	-551906.234464	5.808732e+06	5408.326913	3

Then, we have calculated the distance to nearest bike shop from every area candidate center (not only those within 300m - we want distance to closest one, regardless of how distant it is). And we got the distance data of nearest bike shop, here is shops, shop counts and distance data:

	Address	Latitude	Longitude	X	Y	Distance from center	shops in area	Distance to bike shops
0	89 Stratford Dr, London SW17 0PR	51.434028	-0.178197	-552206.234464	5.808212e+06	5992.495307	0	1116.761952
1	4 John Hunter Ave, London SW17 7FA	51.435132	-0.169867	-551806.234464	5.808212e+06	5840.376700	0	543.775534
2	36 Noyna Rd, Tooting, London SW17 7PH	51.436236	-0.161536	-551006.234464	5.808212e+06	5747.173218	4	253.429288
3	12 Manville Rd, Balham, London SW17 8JN	51.437339	-0.153204	-550406.234464	5.808212e+06	5715.767665	0	726.403860
4	15 Brierley Rd, London SW12 9LY	51.438441	-0.144873	-549806.234464	5.808212e+06	5747.173218	2	856.379581
5	5 Mortimer Cl, St Leonard's, London SW16 1AQ	51.439543	-0.136540	-549206.234464	5.808212e+06	5840.376700	1	1288.064208
6	15 Killieser Ave, Streatham Hill, London SW2 4NX	51.440645	-0.128207	-548606.234464	5.808212e+06	5992.495307	1	1817.919916
7	unit 15, 55-59 Weir Rd, London SW19 8UG	51.436879	-0.192223	-553106.234464	5.808732e+06	5855.766389	0	912.567971
8	49 Waldron Rd, Earlsfield, London SW18 3TA	51.437985	-0.183893	-552506.234464	5.808732e+06	5604.462508	1	1422.743267
9	Clubhouse, Burntwood Ln, London SW17 0AT	51.439089	-0.175562	-551906.234464	5.808732e+06	5408.326913	3	844.235144

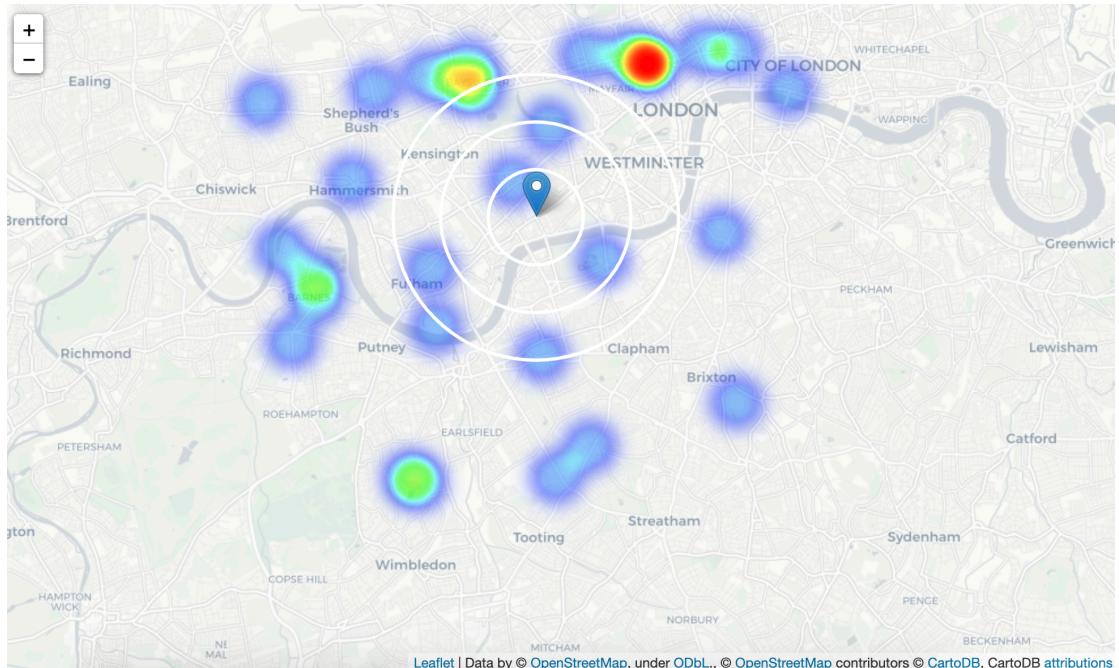
And also, we calculated the Average distance to closest bike shops from each area center: 998.4490421316772, so on average bike shop can be found within ~500m from every area center candidate. That's fairly close, so we need to filter our areas carefully.

We use “folium.Map” to created a map showing heatmap / density of sports shops and try to extract some meaningfull info from that. Also, we have displayed borders of London boroughs on our map and a few circles indicating distance of 1km, 2km and 3km from Chelsea centre. Here is the heat map:



Looks like a few pockets of low shops density closest to city center can be found south, south-east and east from Chelsea centre.

We also have created another heatmap map to showing heatmap/density of bike shops only in Chelsea and its neighborhood. Here is the heatmap:



This map is not so 'hot' (bike shops represent a subset of ~15% of all shops in Chelsea Area London) but it also indicates higher density of existing bike shops directly north and west from Chelsea, with closest pockets of low bike shop 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 London 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 South Kensington and Cremorne Estate which are nearly with Chelsea.

3.2 Analysis of South Kensington and Cremorne Estate

The South Kensington and Cremorne Estate are famous living areas near by Chelsea, and also it has really beautiful view and attract a lot of tourist, it have big business center in the South Kensington and Cremorne Estate. So, South Kensington and Cremorne Estate have a lot of business opportunities.

Analysis of popular travel guides and web sites often mention South Kensington and Cremorne Estate as beautifull, interesting, rich with culture, 'hip' and 'cool' London neighborhoods popular with tourists and loved by London people.

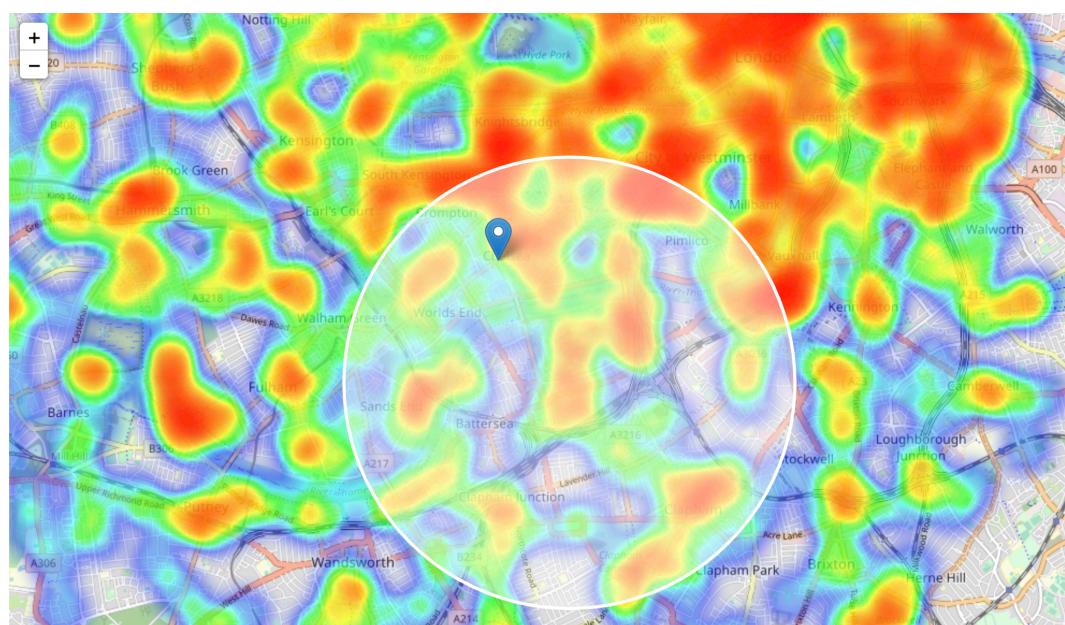
"These spacious, well-equipped self-catering apartments are ideally placed in South Kensington. They offer the chance to be independent and are well suited to families and business travellers; perfect for longer stays too. From the Telegraph.com.

"South Kensington is a very affluent district in West London, just south of Hyde Park. It's home to the Natural History Museum, Science Museum, Imperial College and many many foreign embassies. The area is very popular with the tourists and there man fine cafes and restaurants choose from." From the Tripadvisor.com.

"Possibly one of my favorite, trendy areas in London with a large focus on new and hip (smaller) restaurants with a large variety of international food. You'll be able to find something from every corner of the world and you definitely won't be disappointed! From the trip.com.

So South Kensington and Cremorne Estate are popular with tourists, bikers, alternative and bohemian but booming and trendy, relatively close to underground service, city center and well connected, those boroughs appear to justify further analysis.

We define new, more narrow region of interest, which will include low-shops-count parts of South Kensington and Cremorne Estate closest to Chelsea. Here is the heatmap of shops distribution in South Kensington and Cremorne Estat:



It look not bad - this nicely covers all the pockets of low sport shops density in South Kensington and Cremorne Estate closest to Chelsea.

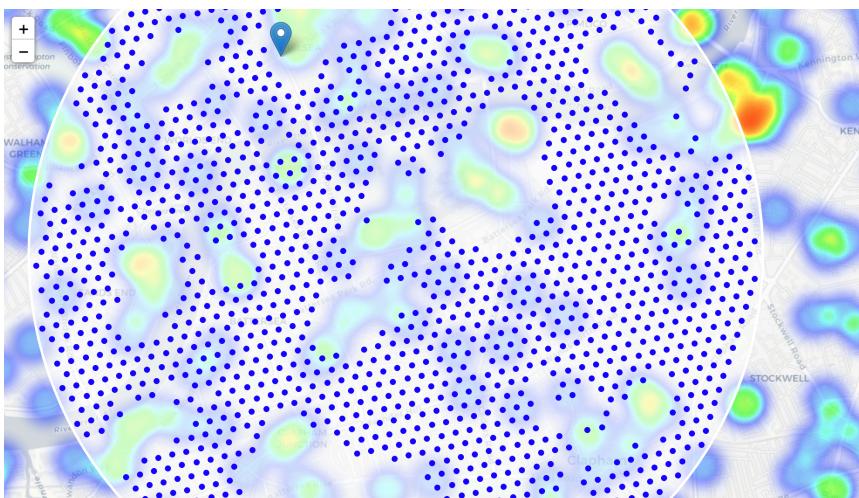
We also create new, more dense grid of location candidates restricted to our new region of interest. We make our location candidates 100m appart. And we have created 2261 candidate neighborhood points. Then, we calculate two most important things for each location candidate: number of sport shops in vicinity (we'll use radius of 250 meters) and distance to closest bike shops. And we save the result in pandas dataframe, here is the data:

	Latitude	Longitude	X	Y	sports nearby	Distance to bike shops
0	51.453057	-0.152001	-549956.234464	5.809928e+06	2	1051.182496
1	51.453241	-0.150612	-549856.234464	5.809928e+06	2	1085.942851
2	51.452797	-0.159896	-550506.234464	5.810014e+06	0	1067.868389
3	51.452981	-0.158507	-550406.234464	5.810014e+06	0	1095.007376
4	51.453165	-0.157118	-550306.234464	5.810014e+06	0	1087.808986
5	51.453349	-0.155729	-550206.234464	5.810014e+06	0	1089.777788
6	51.453533	-0.154340	-550106.234464	5.810014e+06	1	1100.864598
7	51.453717	-0.152951	-550006.234464	5.810014e+06	2	1120.798867
8	51.453900	-0.151562	-549906.234464	5.810014e+06	2	1149.120244
9	51.454084	-0.150173	-549806.234464	5.810014e+06	2	1185.227646

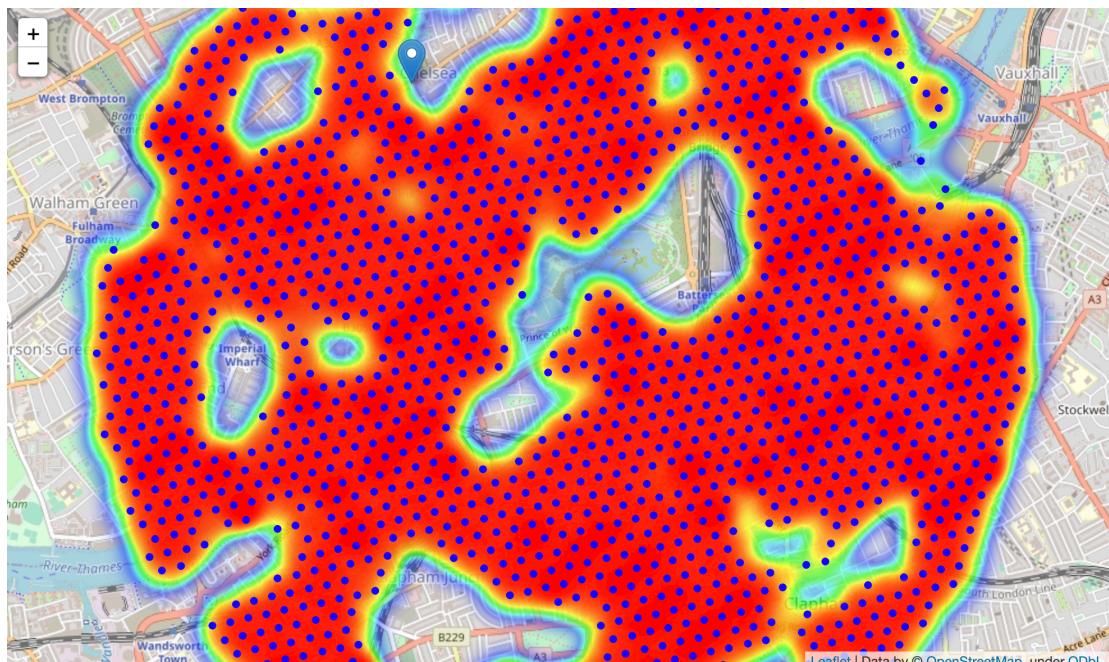
Then

we filter those locations: we're interested only in locations with no more than two shops in radius of 250 meters, and no bike shops in radius of 400 meters. And count the number of sports shops and bike shops, the number of location of only one shop is 1650, number of locations with only one bike shops within 400m is 2114, number of locations with both conditions is 1575.

The we create a map to visualize the locations of sports shops and bike shops in the Chelsea and neighborhood. Here is the map:



Looking good. We now have a bunch of locations fairly close to Chelsea, and we know that each of those locations has no more than two sports shops in radius of 250m, and no bike shops closer than 400m. Any of those locations is a potential candidate for a new bike shop, at least based on nearby competition. Now we show those good locations in a form of heatmap:



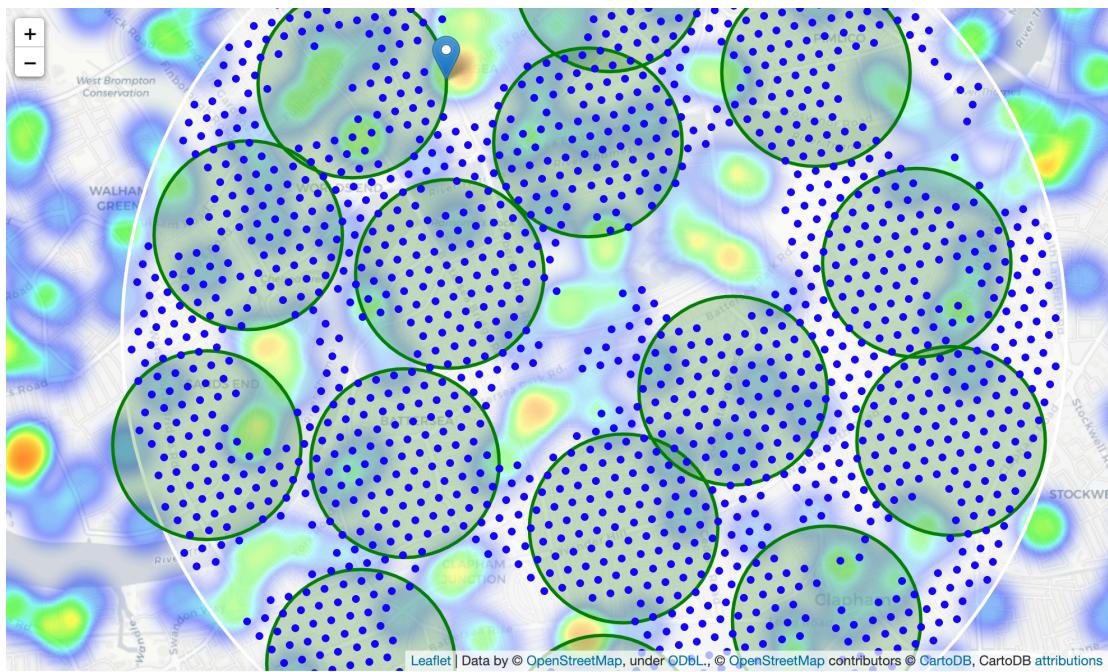
Looking good. What we have now is a clear indication of zones with low number of sports shops in vicinity, and *no* bike shops at all nearby. Now we cluster those locations to create centers of zones containing good locations. Those zones, their centers and addresses will be the final result of our analysis.

4. Cluster Modeling

We use the K-Means Cluster algorithm from Scikit-learn, use cluster algorithm to clustering the sports shops and bike shops, to discover the business opportunities, discover the suitable location for bike shops.

4.1 K-Means model

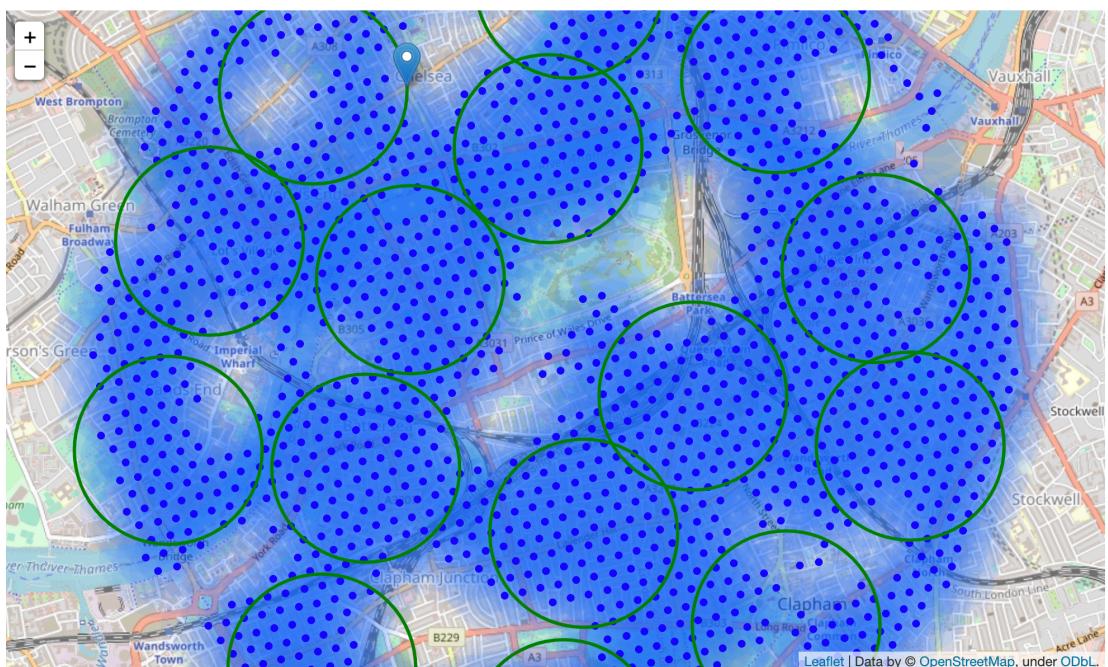
I applied K-Means models, to clustering the sports shops and bike shops, then created the map to display the clustering result, it is clear to show the groups of sports shops and bike shops, the stockholders can select the no or less sports shops and bike shops area to start the bike shop business. Here is the clustering result map:



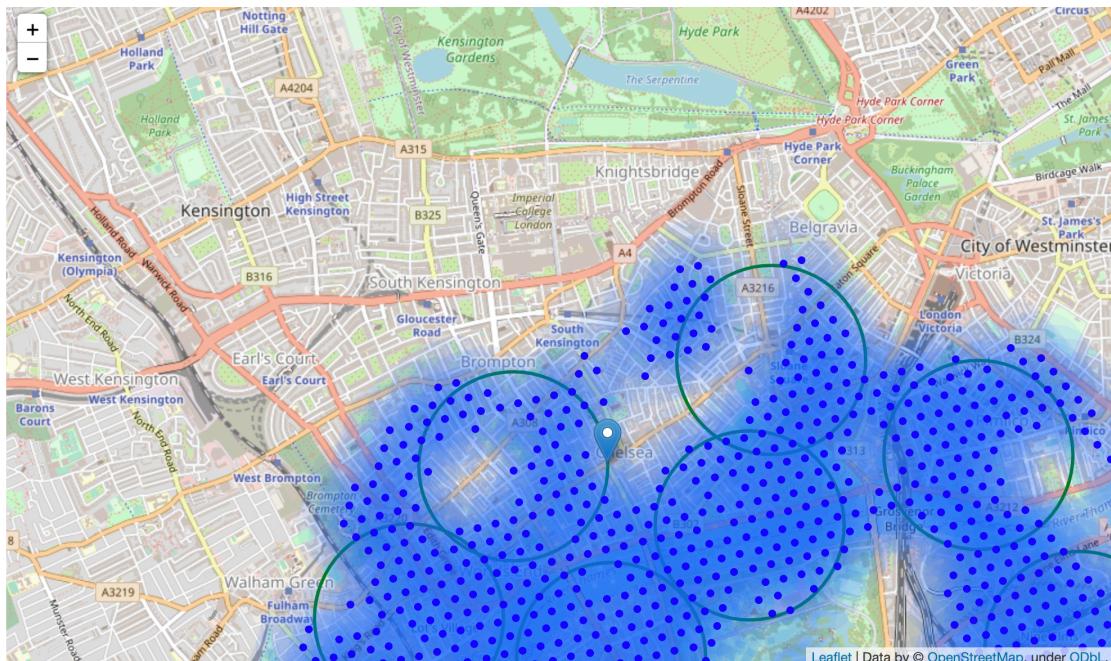
Not bad - our clusters represent groupings of most of the candidate locations and cluster centers are placed nicely in the middle of the zones 'rich' with location candidates.

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.

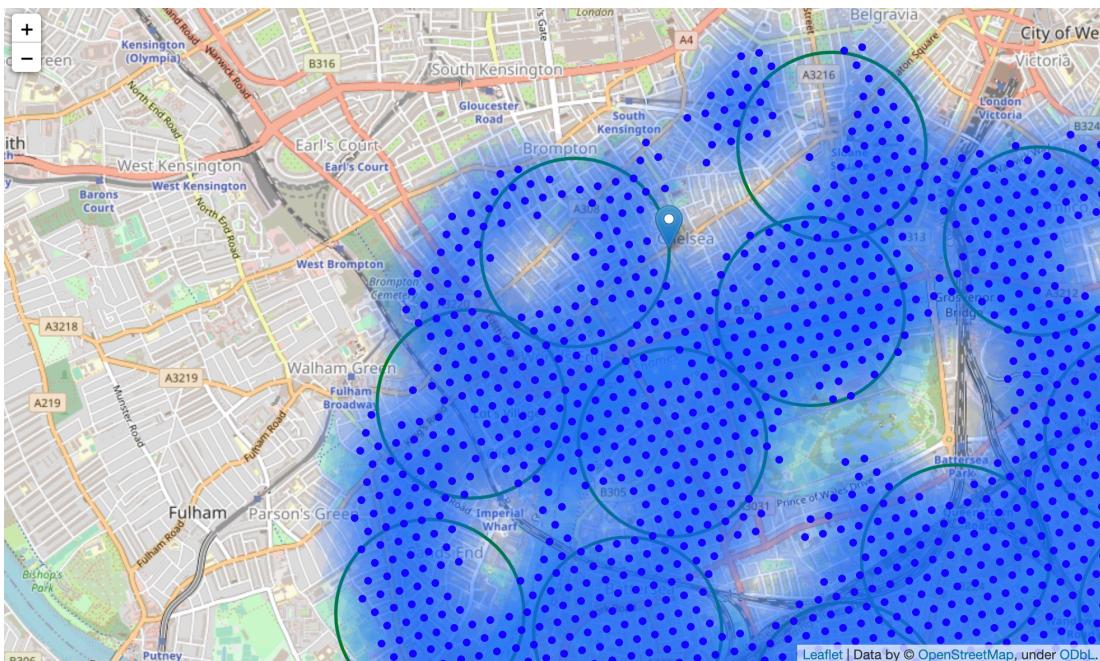
Let's see those zones on a city map without heatmap, using shaded areas to indicate our clusters:



Now zoom in on candidate areas in South Kensington, here is the clustered map of South Kensington:



Also zoom in on candidate areas in Cremorne Estate, here is the clustered map of Cremorne Estate:



Finally, let's reverse geocode those candidate area centers to get the addresses which can be presented to stakeholders. And we got 12 locations are suitable for create the business of bike shops. Here are 12 locations: *Tipthorpe Road (Stop B), Battersea, London*

SW11 5RD => 2.6km from Chelsea centre, 26 Chelsea Embankment, Chelsea, London SW3 4LF=> 0.8km from Chelsea centre, 58 Stephendale Rd, Fulham, London SW6 2PG=> 2.3km from Chelsea centre, 6 St Philip St, London SW8 3SP=> 2.3km from Chelsea centre, 82 Union Rd, Larkhall, London SW4 6JU=> 3.3km from Chelsea centre, 61 Winsham Grove, London SW11 6NE=> 3.6km from Chelsea centre, 134 Fawcett Cl, Battersea, London SW11=> 2.1km from Chelsea centre, Flat 1, 28 Elm Park Gardens, Chelsea, London SW10 9NZ => 0.5km from Chelsea centre, 110 Battersea Bridge Rd, Battersea, London SW11 3AN => 1.0km from Chelsea centre, Langdale House, Pimlico, London SW1V 3DH=> 2.0km from Chelsea centre, 109 Lots Rd, Chelsea, London SW10 0RN=> 1.3km from Chelsea centre, Unnamed Road, Nine Elms, London SW8=> 2.7km from Chelsea centre, 8 Lower Sloane St, Chelsea, London SW1W 8BL=> 1.0km from Chelsea centre, 9 Elsyngue Rd, London SW18 2HW=> 3.2km from Chelsea centre, Long Road Clapham Common, Clapham, London SW4=> 3.5km from Chelsea centre.

This concludes our analysis. We have created 12 addresses representing centers of zones containing locations with low number of sport shop and no bike shop nearby, all zones being fairly close to city center (all less than 4km from Chelsea centre, and about half of those less than 2km from Chelsea). 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 sport shops locations. Most of the zones are located in South Kensington and Cremorne Estate boroughs, which we have identified as interesting due to being popular with tourists, fairly close to city center and well connected by public transport. Here is the map of Chelsea and neighborhood, the blue points represent the suitable bike shop location:



5. Results and Discussion

Our analysis shows that although there is a great number of sport shops in London (~2000 in our initial area of interest which was 12x12km around Chelsea), there are pockets of low sport shop density fairly close to city center. Highest concentration of sport shops was detected north and west from Chelsea, so we focused our attention to areas south, south-east and east, corresponding to boroughs South Kensington, Cremorne Estate and south-east corner of central Mitte borough. Another borough was identified as potentially interesting (Knightsbridge, north-east from Chelsea), but our attention was focused on South Kensington and Cremorne Estate which offer a combination of popularity among tourists, closeness to city center, strong socio-economic dynamics and a number of pockets of low sports shop density.

After directing our attention to this more narrow area of interest (covering approx. 5x5km south-east from Chelsea) we first created a dense grid of location candidates (spaced 100m apart); those locations were then filtered so that those with more than two sport shops in radius of 250m and those with an bike shops 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 12 zones containing largest number of potential new sport shops locations based on number of and distance to existing venues - both sports shop in general and bike shops particularly. This, of course, does not imply that those zones are actually optimal locations for a new sport shop! Purpose of this analysis was to only provide info on areas close to Chelsea center but not crowded with existing sport shop (particularly bike shop) - it is entirely possible that there is a very good reason for medium number of sport shop in any of those areas, reasons which would make them unsuitable for a new sport shop 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.

6. Conclusion

Purpose of this project was to identify London areas close to center with low number of sports shops (particularly bike shops) in order to aid stakeholders in narrowing down the search for optimal location for a new bike shops. By calculating shops density distribution from Foursquare data we have first identified general boroughs that justify further analysis (South Kensington and Cremorne Estate), and then generated extensive collection of locations which satisfy some basic requirements regarding existing nearby shops. 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 decision on optimal sports shop 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.