Predicting the optimal location for a bike shop in London

Predicting the optimal location for a bike shop in London

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.

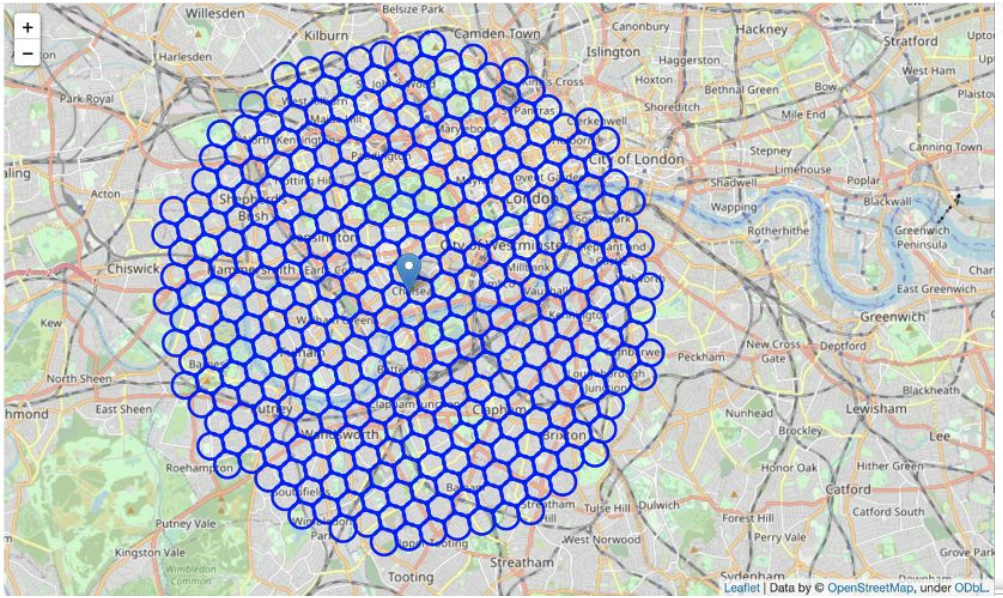
Data acquisition and cleaning

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 aprox. 12x12 killometers centered around London Chelsea center. Find the latitude & longitude of London city center, using specific, well known address and Google Maps geocoding API.

Create a grid of area candidates, equaly 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.

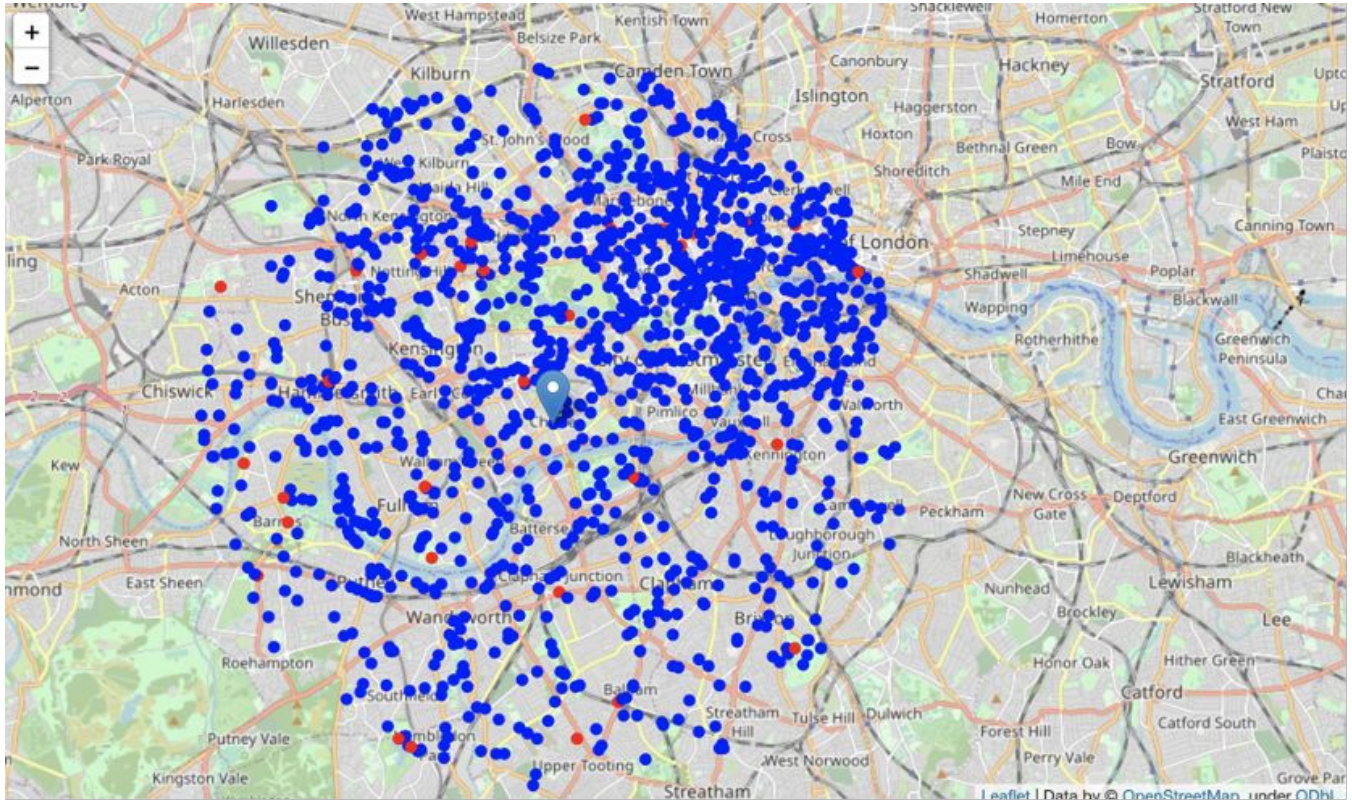
Got coordinate of Chelsea, London, UK: [51.486943, -0.170037]. Then create a hexagonal grid of cells: we offset every other row, and adjust vertical row spacing so that every cell center is equally distant from all it's neighbors.

Created 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:




```
['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',
```

Using 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:



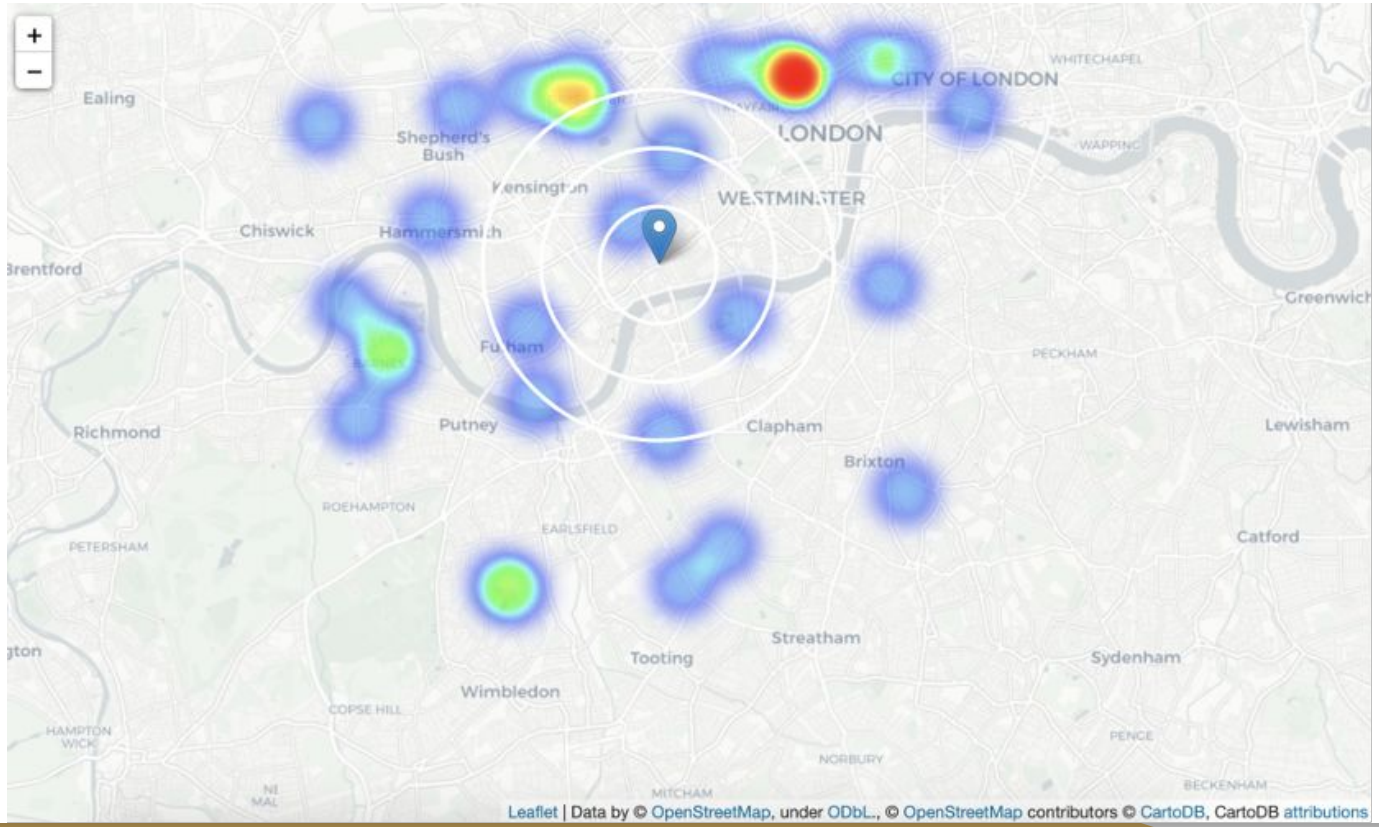
Exploratory Data Analysis

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, and we got the distance data of nearest bike shop, here is the data:

	Address	Latitude	Longitude	X	Y	Distance from center	shops in area	Distance to bike shops
0	89 Strathdon 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	-551606.234464	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 La, London SW17 0AT	51.439089	-0.175562	-551906.234464	5.808732e+06	5408.326913	3	844.235144

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:

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

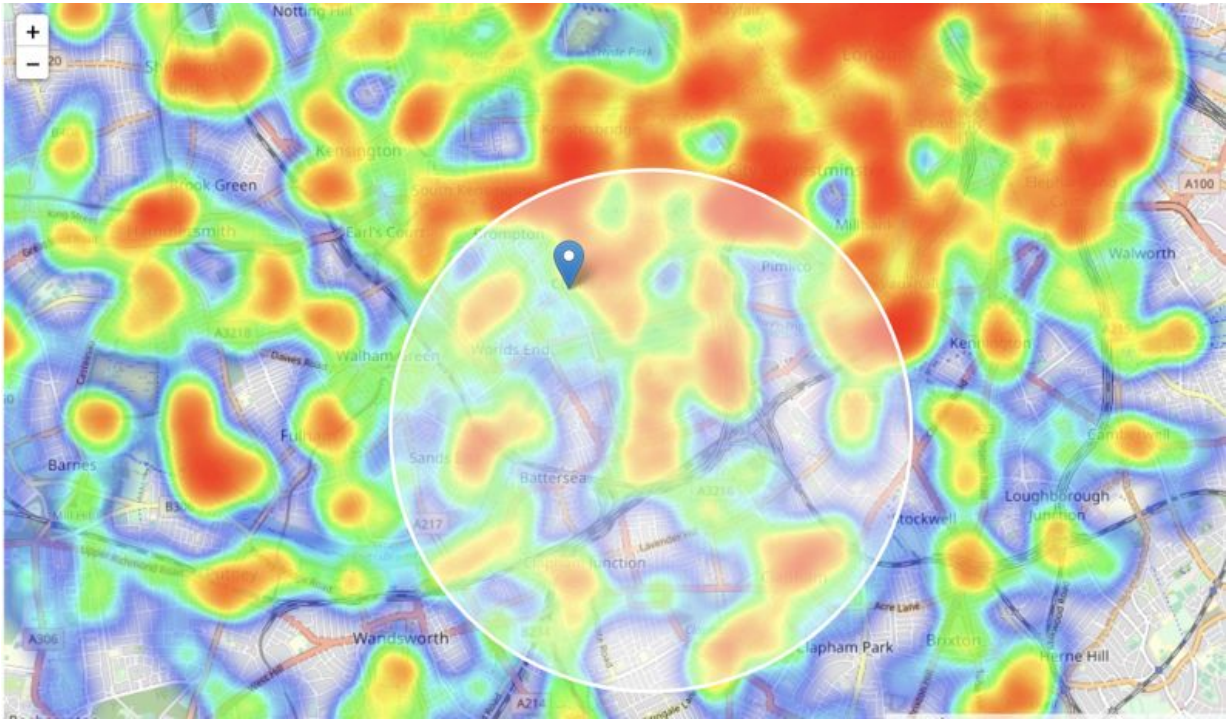


Analysis of South Kensington and Cremorne Estate

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.

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 Estate:

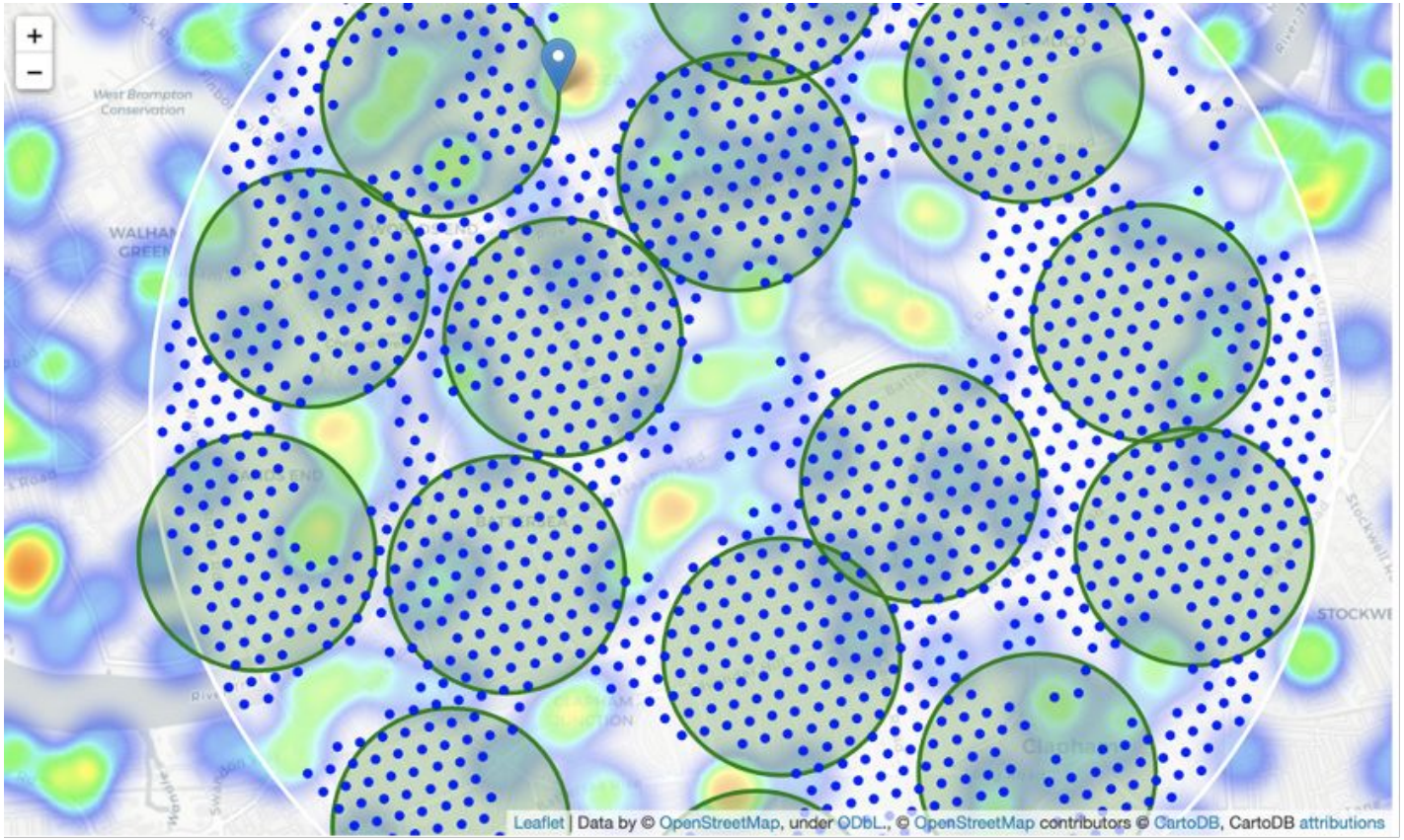


Cluster Modeling

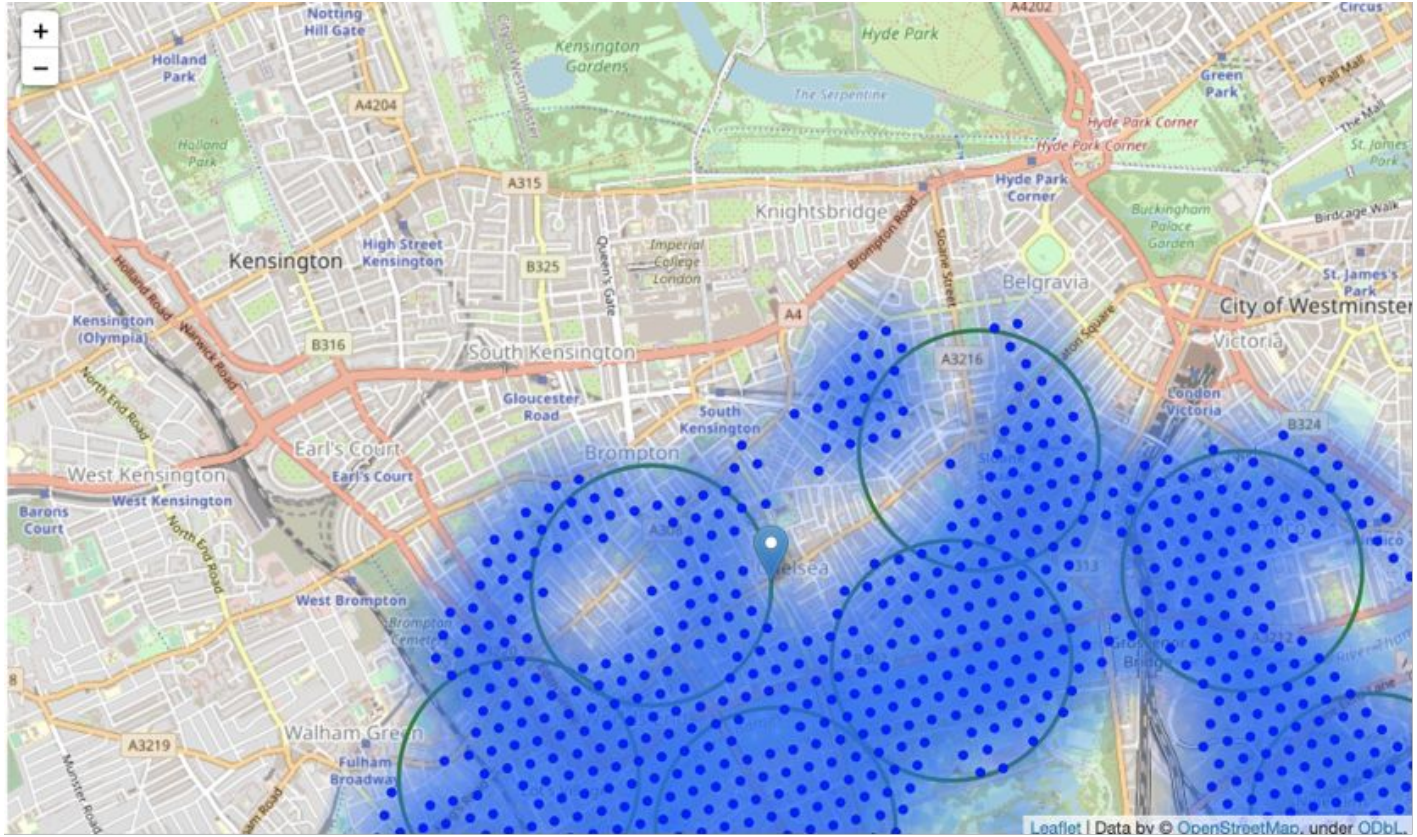
We use the K-Means Cluster algorithm form Scikit-learn, use cluster algorithm to clustering the sports shops and bike shops, to discover the business opportunities, discover the suitable location for bike shops.

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.

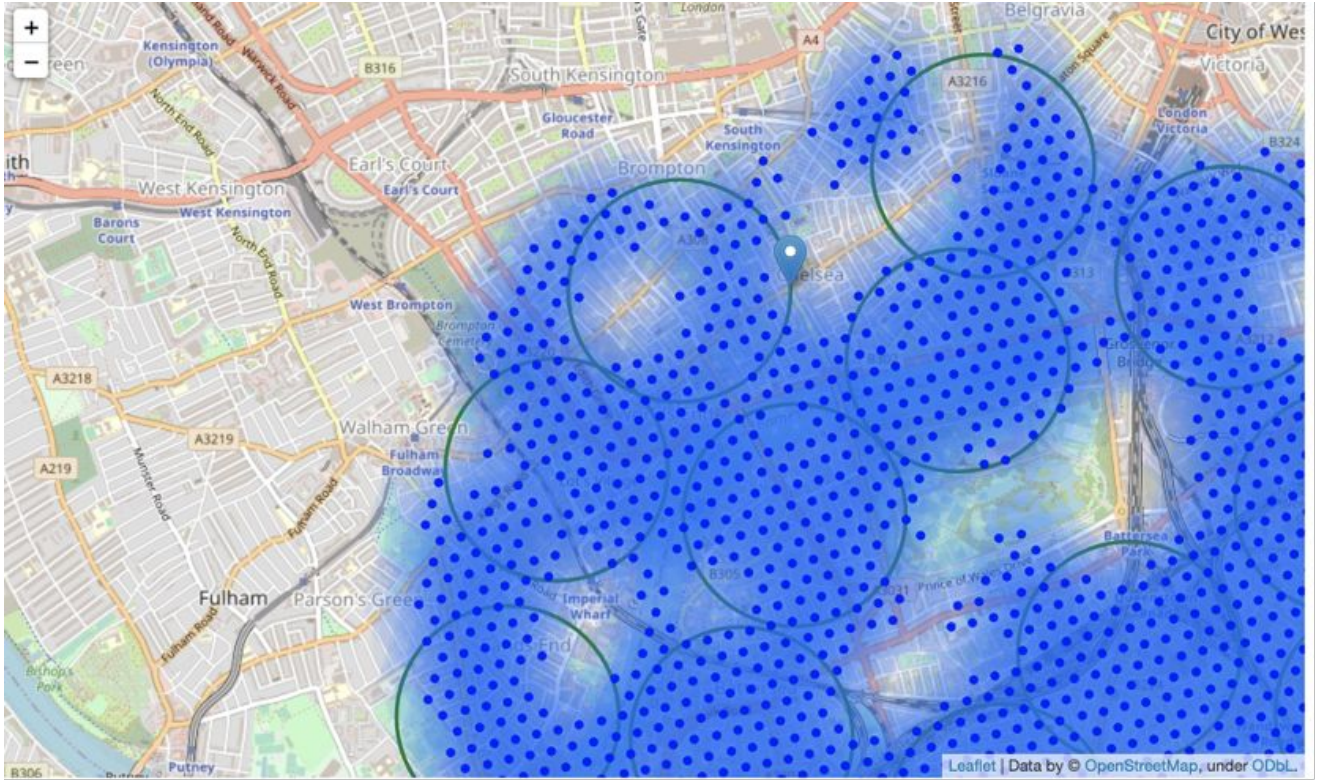
Here is the clustering result map:



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:



Here is the map of Chelsea and neighborhood, the blue points represent the suitable bike shop location:

Results and Discussion

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.

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.