London Venues Analysis

February 19th, 2021

By Marco Andrea MAGGI

Table of Contents

Business Problem

Our customer - a medium-sized company operating in the coffee shop retailing - is looking for suited boroughs in the **Greater London Area** to open three new coffee shops plus one flagship store.

The main requirements regarding these "one-plus-three" shops are listed below:

1.   The shops must be well spread in the Greater London Area.
2.   The shops should be placed in areas with low concentration of competitors: **"competitors"** are those franchises featuring a number of shops in the same area ranging from 3 to 6.
3.   The shops should be placed in areas with proven Business Resilience (i.e. relatively high business survival rate).
4.   The shops should be placed in mid\high populated areas (i.e. mid\high population density).
5.   The flagship shop must be placed in one of the **INNER LONDON** boroughs.
6.   The other three shops must be placed - each of them - in a suited borough in the **OUTER LONDON** boroughs (provided with the features described in the previous points) chosen among the boroughs belonging to three geographical areas to be found.

Our company has been commissioned to prepare a detailed analysis regarding suitable places to open these one plus three shops, pointing out the strong assets we find out by taking into account the requirements.

Data Source

The source of data used in this analysis are:

- The **London Datastore (London Boroughs Profile)** by Greater London Authority (License)
- **Foursquare** location platform

The **London Boroughs Profile** - managed by Greater London Authority - is meant to (quoting):

"... help paint a general picture of an area by presenting a range of headline indicator data in both spreadsheet and map form to help show statistics covering demographic, economic, social and environmental datasets for each borough, alongside relevant comparator areas."

In this framework, the London Boroughs Profile is helpful to provide the needed information to fulfill the requirements, information used was updated during 2020 (more info about updating time unavailable).

Foursquare is a world-famous provider for BI analysis and data analytics meant to provide (quoting):

"... Information about people and information about places - we package them together into practical and flexible data sets that you can add in your BI tools and customer databases. Urban planning, financial modeling, and customer segmentation are just a few areas where an understanding of human movement will lead to smarter outcomes. "

Foursquare will be used to explore London boroughs in search of existing shops.

Information collected from these two sources will be combined to build up the final picture.

Analysis Methodology and Results

A subset of all data gathered from London Boroughs Profile will be loaded, the core information collected is:

1. **Borough Code** ['Code']
2. **Borough Name** ['Area_name']
3. **Inner or Outer London Id** ['Inner/_Outer_London']
4. **Land Area** ['Inland_Area_(Hectares)']
5. **Population Density** ['Population_density_(per_hectare)_2017']
6. **Average Age** ['Average_Age,_2017']
7. **Working Age Percentage** ['Proportion_of_population_of_working-age,_2015']
8. **Business Resilience** ['Two-year_business_survival_rates_(started_in_2013)']

Some data wrangling is needed to remove extra-rows featuring the same eight columns listed above and regarding the whole London Area, England and United Kingdom stats.

Information 1-5 and 8 will be used in this paper (marked in yellow), the remaining ones can be used to further deepen the borough characterization.

As a general comment, this dataset is already cleaned, no missing values are present.

The **Greater London Area** has **33 boroughs**: they are shown below - Fig.1 - and they belong to **Inner London** or **Outer London**.
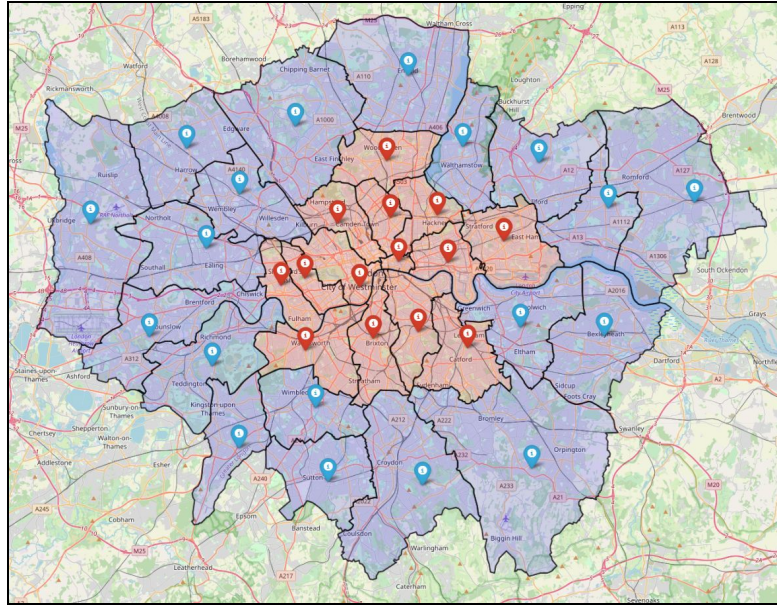
Fig. 1 - London Boroughs and their geographical references (the black dots)

Thanks to the OpenStreetMapTool Python package, we can retrieve the geographical coordinates of each borough center (shown in Fig. 2): each center is shown in the map and it is used to **CLUSTERIZE** the **Outer London Area.**
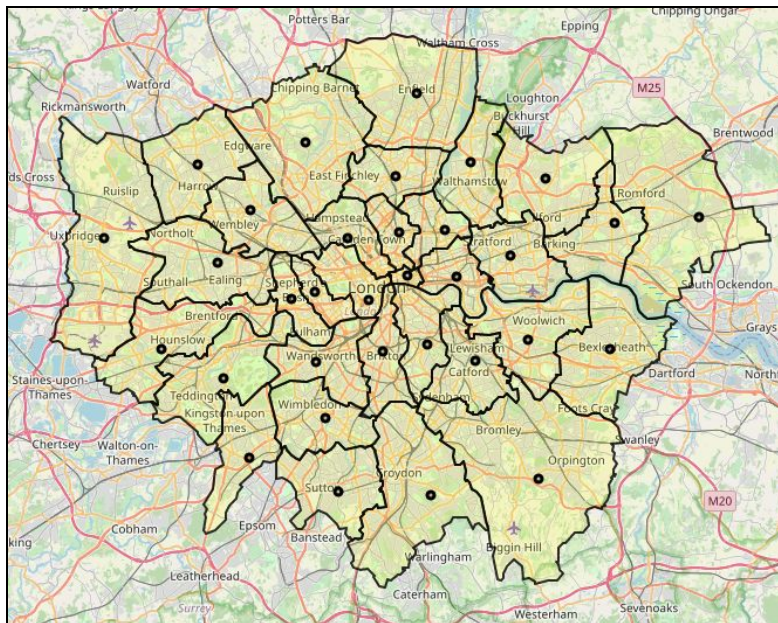

Fig. 2 - London Boroughs and their geographical references (the black dots)

Three different clusters are created, the three "standard" shops will be placed in a borough belonging to one cluster (i.e. one shop per one cluster).

"Clusterization" will be achieved by applying **K-Means** algorithm to geographical coordinates of each borough: **this way of clustering the boroughs is meant to divide the borough per geographical areas**.

The clusters - in blue - are shown in Fig. 3: the red area is featuring the **INNER LONDON** area.
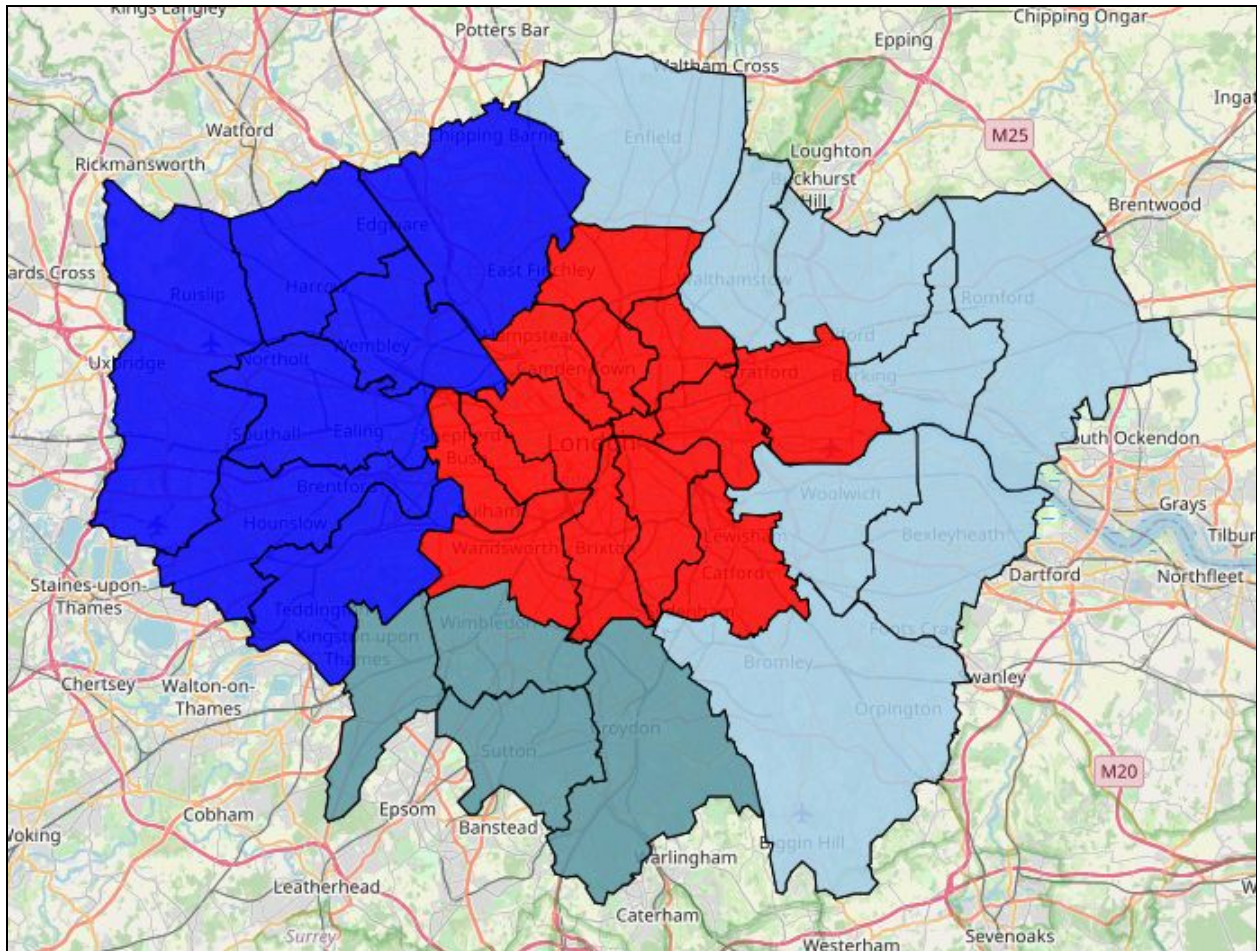

Fig. 3 - London Boroughs and their clustering based on geographical coordinates

The core characteristics of each cluster are shown in the Table 1.

Area	Boroughs No	Cluster No	Color
INNER LONDON	14	#0	Red
OUTER LONDON/WEST AREA	7	#1	Dark Blue
OUTER LONDON/NORTH AND EAST AREA	8	#2	Light Blue
OUTER LONDON/SOUTH AREA	4	#3	Cadet Blue

Table. 1 - Cluster Resume

The 33 boroughs have been resumed in four geographical areas:

- **INNER London - Center Area**
- **OUTER London - West Area**
- **OUTER London - North and East Area**
- **OUTER London - South Area**

Each area (or cluster) has a **different number of boroughs.**

Now we are ready to retrieve information regarding the **coffee shops** in the London Area: Foursquare allows us to collect information in the neighborhood of a geographical point (circle with a 3.5 km radius). A sample of available data - featuring the first and the last five- is shown in Fig. 4

	Area_name	Venue	Venue Latitude	Venue Longitude	Venue Category
0	City of London	WA Cafe	51.51101	-0.12661	Café
1	City of London	Monmouth Coffee Company	51.51431	-0.12682	Coffee Shop
2	City of London	Coffee Island	51.51245	-0.12718	Coffee Shop
3	City of London	Lundenwic	51.51282	-0.11834	Coffee Shop
4	City of London	% Arabica	51.51173	-0.12405	Coffee Shop
...	...	...	...	...	...
976	Westminster	Hagen	51.50982	-0.13781	Coffee Shop
977	Westminster	The Roasting	51.49136	-0.13896	Coffee Shop
978	Westminster	Prufrock Coffee	51.51993	-0.10945	Coffee Shop
979	Westminster	L'ETO Caffè	51.51434	-0.13467	Café
980	Westminster	Roasting Plant Coffee	51.50626	-0.08835	Coffee Shop

Fig. 4 - Coffee shops in London by Foursquare

981 shops have been found: their distribution over the London area is not homogeneous.

It is worth to notice that "three big players" (Costa Coffee\Starbucks\Caffè Nero) are present in the market, **summing up the 21.7% of the total number of shops in the area (Fig. 5)**.
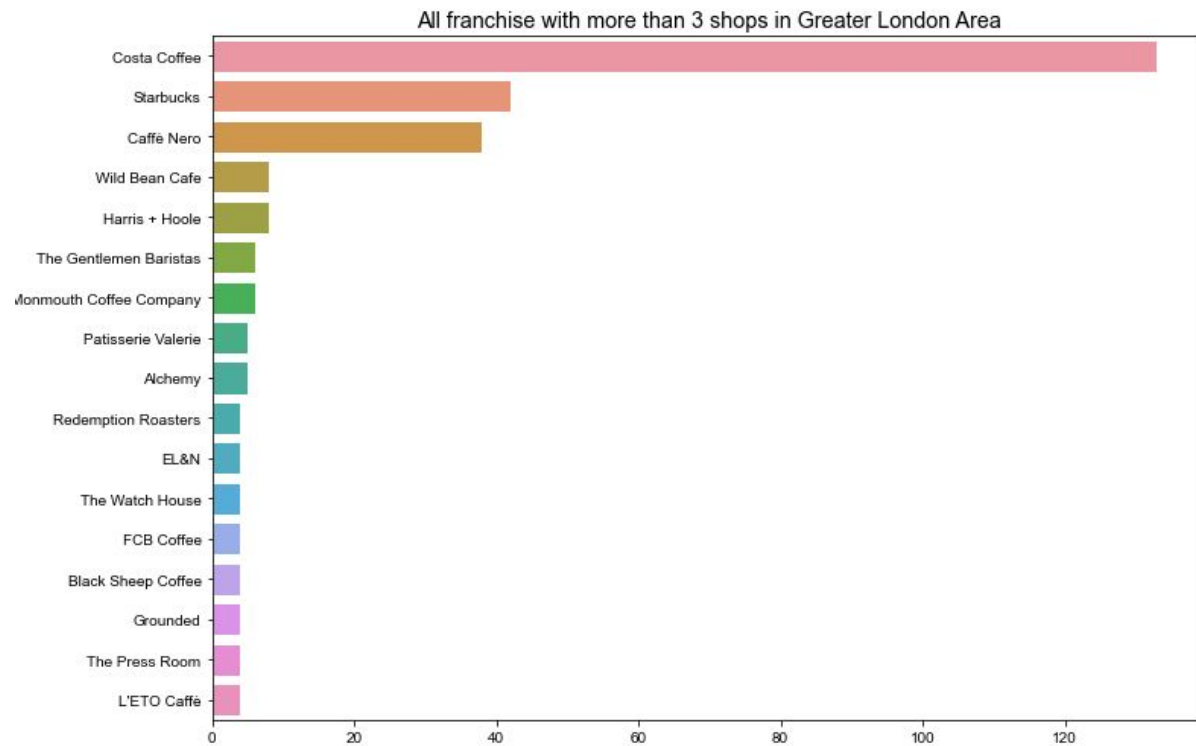
Fig. 5 - The number of shops of those retailer with more than three shops in London

However if we restrict our analysis to the "competitors" of our customer (i.e. franchise with a number of shops ranging from 3 to 6) **we have to restrict our analysis to 129 total shops.**

Pairing the general stats and clustering with venue information, we can build up a snapshot of the borough general stats (demographic+economic) and the presence of competitor shops in the borough area.

The next figure (Fig. 6) resumes the distribution in terms of competitors shops per borough.
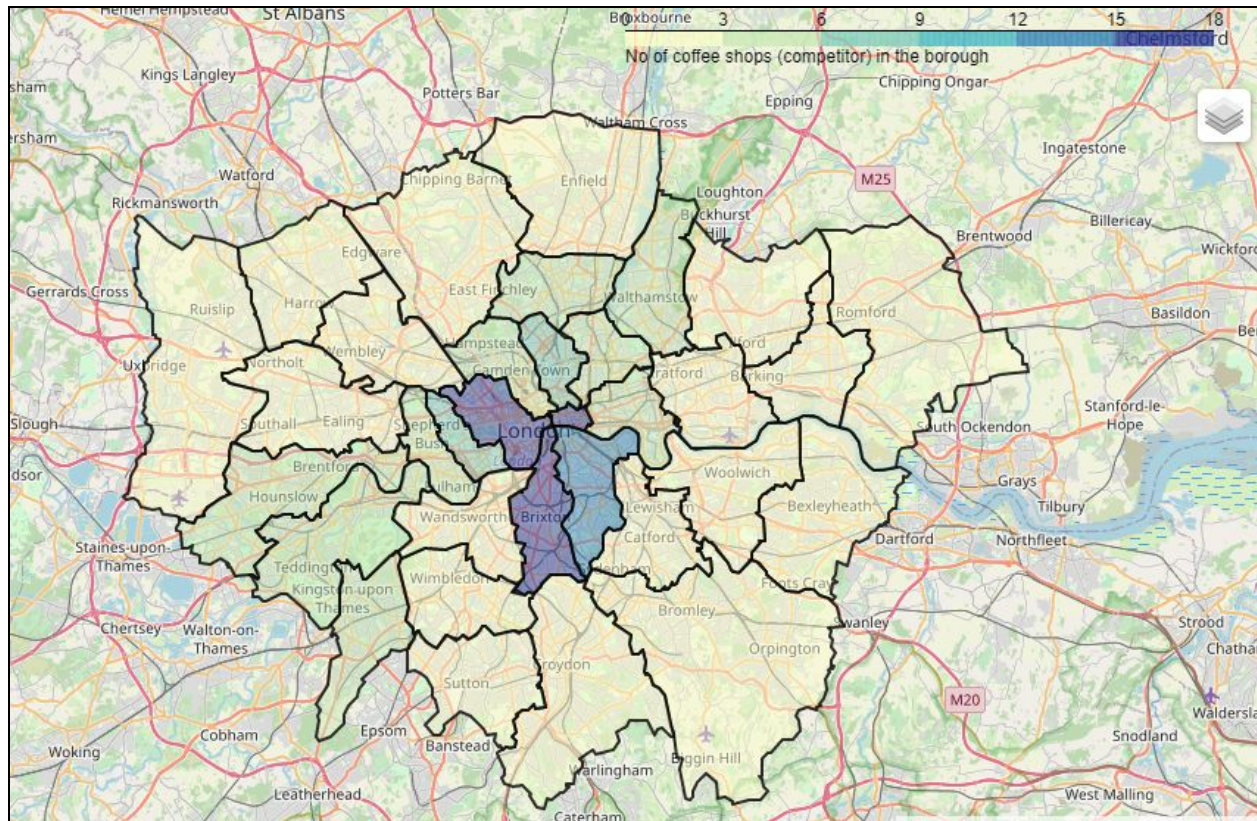
Fig. 6 - The number of competitors shops in London

It is clear that some areas in the INNER LONDON have a sharp concentration of shops.

Following the business problem description, we are called to define three index per each borough, featuring

1. **The density of competitors coffee shops in the area**
2. **The business survival index**
3. **The population density**

1 is meant as **"the lower, the better"**, **2** and **3** are meant as **"the higher, the better".**

The number of competitor shops per unit of inland area is used to assess 1. (by assigning "0" to the most crowded area and "1" to the less crowded area and linearly mapping the intermediate), **the business survival index by Greater London Authority is used to assess 2.** (by

normalizing the index with max value of this index), **the population density by Greater London Authority to assess 3.** (again, by normalizing with max value of pop density).

The three indexes are ranging between 0 and 1 and they can be linearly combined, in this way, to create a final Area Appeal Index: **the higher is that index, the better is the area having that index to place a new shop.**

The Table 2 is showing the coefficients used to combine the three indexes to get the appeal index.

FACTORS	WEIGHTS
CoffeeShop Density per Ha	45%
Business Survival Index	25%
Population Density	30%

Table 2 - Weighting Factors
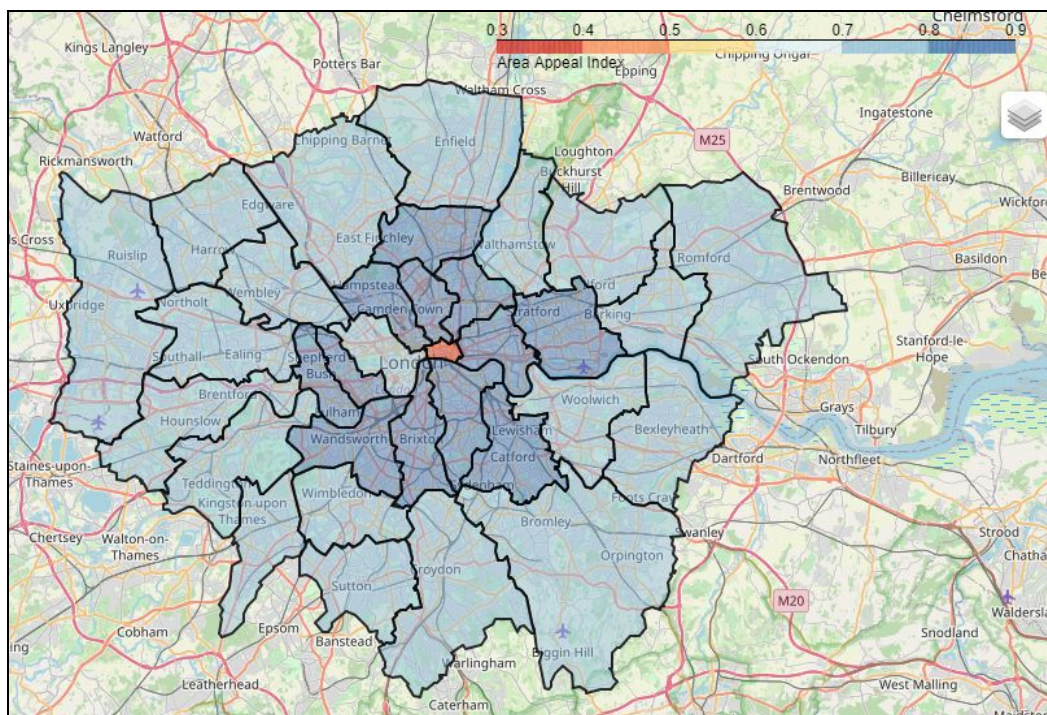
The final ranking is resumed in the Fig. 7

Fig. 7 The boroughs ranking by Area Appeal Index

The Table 3 resumes all the Area Appeal Index, divided by Clusters (0=INNER London, 1-2-3 are the OUTER London clusters defined in the previous pages), **the best performing area per cluster is also highlighted**.

| PIVOT TABLE | Cluster No | | | |
Area Name	0	1	2	3
Barking and Dagenham			0.791	
Barnet		0.771		
Bexley			0.761	
Brent		**0.832**		
Bromley			0.740	
Camden	0.887			
City of London	0.262			
Croydon				0.774
Ealing		0.810		
Enfield			0.763	
Greenwich			0.791	
Hackney	**0.950**			
Hammersmith and Fulham	0.886			
Haringey	0.853			
Harrow		0.789		
Havering			0.732	
Hillingdon		0.738		
Hounslow		0.782		
Islington	0.947			
Kensington and Chelsea	0.899			
Kingston upon Thames				0.776
Lambeth	0.843			
Lewisham	0.845			
Merton				**0.805**
Newham	0.852			
Redbridge			0.789	
Richmond upon Thames		0.759		
Southwark	0.859			
Sutton				0.776
Tower Hamlets	0.947			
Waltham Forest			**0.802**	
Wandsworth	0.871			
Westminster	0.817			

Table 3 - Area Appeal Index and the best performer per each cluster

Discussion

The most attractive areas are in **Inner London** (the first 12 boroughs are in that area), **West Outer London** looks a little bit less attractive on average than the other two areas.

Let's have a deep dive into the results:

- *Cluster #0 (Inner London)*

  Hackney is the best borough to place our flagship store, the main metrics for all the borough in Inner London Area are listed in Table 4

Cluster #0 Ranking:

	Area_name	Extended_name	Lat	Lon	ClusterArea	NoOfShops	Shop_Density_per_Hectar	Business_Survival	Population_Density	Area_Appeal_Index
11	Hackney	London Borough of Hackney, London, Greater Lon...	51.54888	-0.04767	0	5	0.95243	0.97462	0.92545	0.94988
18	Islington	London Borough of Islington, London, Greater L...	51.54703	-0.10166	0	6	0.92682	0.92005	1.00000	0.94708
29	Tower Hamlets	London Borough of Tower Hamlets, London, Great...	51.51456	-0.03501	0	5	0.95418	0.88452	0.98779	0.94685
19	Kensington and Chelsea	Royal Borough of Kensington and Chelsea, Londo...	51.50380	-0.20079	0	6	0.91027	0.94543	0.84254	0.89874
6	Camden	London Borough of Camden, London, Greater Lond...	51.54285	-0.16251	0	3	0.97505	0.93401	0.71530	0.88686
12	Hammersmith and Fulham	London Borough of Hammersmith and Fulham, Lond...	51.49831	-0.22788	0	3	0.96684	0.93147	0.72622	0.88581
31	Wandsworth	London Borough of Wandsworth, London, Greater ...	51.45190	-0.19951	0	0	1.00000	0.96193	0.60219	0.87114
27	Southwark	London Borough of Southwark, London, Greater L...	51.46528	-0.06904	0	12	0.92464	0.93147	0.69987	0.85892
13	Haringey	London Borough of Haringey, London, Greater Lo...	51.58793	-0.10541	0	5	0.96938	0.94416	0.60347	0.85330
24	Newham	London Borough of Newham, London, Greater Lond...	51.53000	0.02932	0	1	0.99499	0.88832	0.60861	0.85241
22	Lewisham	London Borough of Lewisham, London, Greater Lo...	51.45343	-0.01251	0	2	0.98969	0.93147	0.55463	0.84462
21	Lambeth	London Borough of Lambeth, London, Greater Lon...	51.46040	-0.12135	0	15	0.89859	0.80964	0.78856	0.84335
32	Westminster	City of Westminster, London, Greater London, E...	51.49732	-0.13715	0	18	0.84819	0.87310	0.72429	0.81725
0	City of London	City of London, Greater London, England, Unite...	51.51562	-0.09200	0	16	0.00000	0.81599	0.19473	0.26242

Table 4 - Inner London Borough Ranking

The strong point in Hackney is about Business Survival Index, with mid-high density of competitor shops and high population density: particular mention goes to the Business Survival Index, showing a very good business resilience in that area.

It is worth noting the poor performance of the City of London, mostly driven by very high presence of competitors paired with business volatility and low population density: honestly, Shop Density Index clearly is over-penalizing that area.

Average Index in this area is 0.837.

- *Cluster #1 (West Outer London)*

  Waltham Forest wins in the West Area of Outer London, thanks to a pretty low density of competitor shops paired with good Business Resilience, but, a special mention goes to Population Density, which, despite this index is generally small in comparison with same index in Cluster #0, it is sharply higher in Waltham Forest in comparison with other boroughs in this area (see Table 5).

Cluster #1 Ranking:

	Area_name	Extended_name	Lat	Lon	ClusterArea	NoOfShops	Shop_Density_per_Hectar	Business_Survival	Population_Density	Area_Appeal_Index
30	Waltham Forest	London Borough of Waltham Forest, London, Grea...	51.59817	-0.01784	1	5	0.97665	0.90102	0.45758	0.80202
10	Greenwich	Royal Borough of Greenwich, London, Greater Lo...	51.46863	0.04884	1	2	0.99234	0.92259	0.38046	0.79134
1	Barking and Dagenham	London Borough of Barking and Dagenham, London...	51.55412	0.15050	1	1	0.99498	0.92640	0.37211	0.79097
25	Redbridge	London Borough of Redbridge, London, Greater L...	51.58637	0.06976	1	1	0.99679	0.94797	0.34640	0.78947
9	Enfield	London Borough of Enfield, London, Greater Lon...	51.64874	-0.08098	1	2	0.99552	0.94162	0.26478	0.76282
3	Bexley	London Borough of Bexley, London, Greater Lond...	51.46197	0.14570	1	0	1.00000	0.93274	0.25900	0.76088
5	Bromley	London Borough of Bromley, London, Greater Lon...	51.36686	0.06171	1	2	0.99759	0.99746	0.14010	0.74031
15	Havering	London Borough of Havering, London, Greater Lo...	51.55793	0.24981	1	1	0.99839	0.95558	0.14524	0.73174

Table 5 -West Outer London Borough Ranking

Average Index in this area is 0.771.

- *Cluster #2 (North and East Outer London)*

  Brent wins in the North and East Area of Outer London, featuring a very low density of competitor shops, a pretty good business resilience and a good population density (the situation of this borough in its area is similar to Waltham Forest), see Table 6.

Average Index in this area is 0.783.

Cluster #2 Ranking:

	Area_name	Extended_name	Lat	Lon	ClusterArea	NoOfShops	Shop_Density_per_Hectar	Business_Survival	Population_Density	Area_Appeal_Index
4	Brent	London Borough of Brent, London, Greater Londo...	51.56383	-0.27576	2	1	0.99581	0.94416	0.49357	0.83223
8	Ealing	London Borough of Ealing, London, Greater Lond...	51.52508	-0.31429	2	2	0.99347	0.96193	0.40681	0.80959
14	Harrow	London Borough of Harrow, London, Greater Lond...	51.59683	-0.33732	2	0	1.00000	0.97081	0.32134	0.78910
17	Hounslow	London Borough of Hounslow, London, Greater Lo...	51.46173	-0.38013	2	3	0.99029	0.96701	0.31491	0.78185
2	Barnet	London Borough of Barnet, London, Greater Lond...	51.61252	-0.21144	2	0	1.00000	0.93655	0.28856	0.77071
26	Richmond upon Thames	London Borough of Richmond upon Thames, London...	51.44055	-0.30764	2	5	0.98421	1.00000	0.22108	0.75922
16	Hillingdon	London Borough of Hillingdon, London, Greater ...	51.54252	-0.44833	2	0	1.00000	0.95178	0.16710	0.73807

Table 6 - North and East Outer London Borough Ranking

- ***Cluster #3 (South Outer London)***

    <mark>Merton</mark> wins in the South Area of Outer London, featuring no competitor shops, a really good business resilience but it is penalized by very low population density, see Table 7.

	Area_name	Extended_name	Lat	Lon	ClusterArea	NoOfShops	Shop_Density_per_Hectar	Business_Survival	Population_Density	Area_Appeal_Index
23	Merton	London Borough of Merton, London, Greater Lond...	51.41087	-0.18810	3	0	1.00000	0.99492	0.35540	0.80535
28	Sutton	London Borough of Sutton, London, Greater Lond...	51.35746	-0.17363	3	2	0.99173	0.96447	0.29692	0.77647
20	Kingston upon Thames	Royal Borough of Kingston upon Thames, London,...	51.38178	-0.27699	3	4	0.98054	0.97462	0.30270	0.77571
7	Croydon	London Borough of Croydon, London, Greater Lon...	51.35506	-0.06431	3	1	0.99790	0.95558	0.28728	0.77414

Table 7 - South Outer London Borough Ranking

Average Index in this area is 0.783.

Conclusions

So, just to recap, a detailed analysis of the Greater London Area has been done, by collecting some metrics to have a picture about the demographic and general economic environment.

This information has been merged with on-purpose information regarding coffee shops to draw a comprehensive description of Greater London Market: all of these info have concurred in the detection of four boroughs to open one flagship (in <mark>Hackney</mark>) plus three new standard shops (in <mark>Waltham Forest</mark>, <mark>Brent</mark> and <mark>Merton</mark>).