

# London Real Estate Analytics

Fuad Hassar

July 28, 2019

## 1. Introduction

### 1.1 Background

The city of London is the capital of UK and most populous city in UK. It provides lot of business opportunities and business friendly environment. It has attracted many people for living in London. It is a diverse and main tourist city in UK. It is a global hub of the world. The city is a major center for tourism, real estate, entertainment, theater, fashion, and the arts in the United Kingdom. This also means that the market of real estate business is highly competitive. As it is highly developed city, buying house is one of the highest cost. Thus, buying house in this fabulous city needs to be analyzed carefully. The insights derived from analysis will give good understanding of the proper cost of houses in London and will help us for clear strategically decision. Also, guide in reduction of risk and the return on investment to be reasonable.

### 1.2 Problem

In this scenario, it would be a great to adopt machine learning tools to assist homebuyers to make wise and effective decisions. As a result, the business problem that proposing is how could we provide support to homebuyers clientele to purchase a suitable real estate in London? in this uncertain economic and financial scenario?

To solve this business problem, we are going to cluster London neighborhoods in order to recommend venues and the current average price of real estate where homebuyers can make a real estate investment. We will recommend profitable venues according to amenities and essential facilities surrounding such venues, example elementary schools, high schools, hospitals & grocery stores.

### 1.3 Interest

Obviously, real estate agencies and investors who have interest in the investments in London or homebuyers would be very interested in accurate prediction of reasonable houses prices to help guide in reduction of risk and the return on investment.

## **2. Data acquisition and cleaning**

### **2.1 Data sources**

Data related to London properties including houses prices extracted from the Land Registry (<http://landregistry.data.gov.uk/>). The following fields comprise the address data included in Price paid Data: Postcode, Primary Addressable, Object Name. Street; Town; District.

The Foursquare API will be used to explore locations across different venues according to the presence of amenities and essential facilities. The Foursquare explore function will be used to get the most common venue categories in each neighborhood, and then use this feature to group the neighborhoods into clusters. The k-means clustering algorithm will be used for the analysis. In the end, the Folium library will be used to visualize locations and facilities and their emerging clusters. In the end, recommendation of profitable real estate investments will be given.

### **2.2 Data cleaning**

Data downloaded or scraped from mainly two sources were data got massaged and stored in different data frames. The data is very huge where decided to use data from 2017 until 2019. The Data columns got renamed and formatted to be suitable for machine learning algorithm according to the solution scope. Data got filtered based on city of London , list of street names in London, Average prices set to range between 2000000 to 25000000. In addition, Calculate the street-wise average price of the property and read the street-wise coordinates into a data frame. The data got Joined to find the coordinates of locations and according to clients budgets. Finally, visualization of map plot recommended locations on London map addressing all venues.

Table 1. Data Consolidation

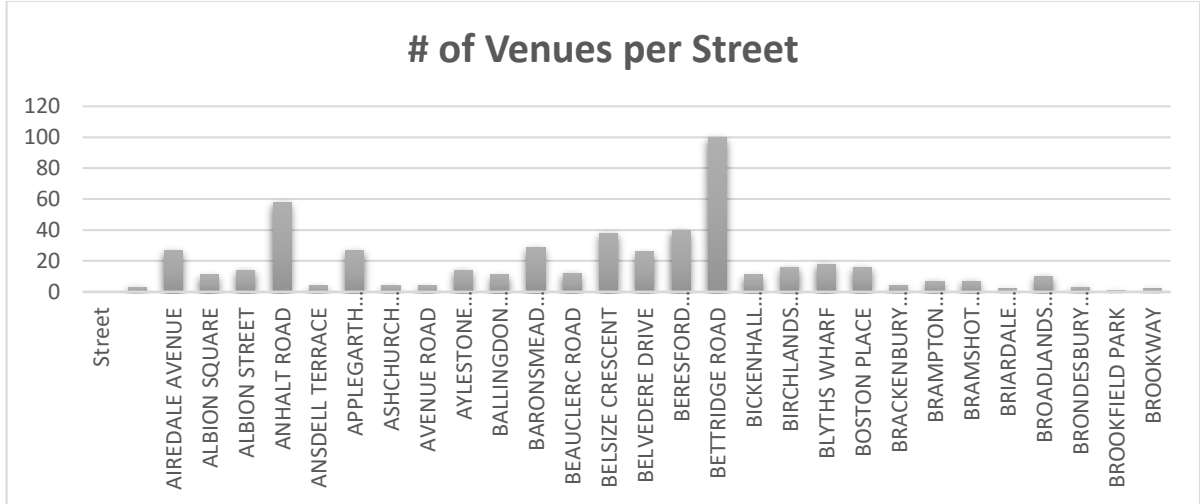
	Street	Avg_Price	city_coord	Latitude	Longitude
146	AIREDALE AVENUE	2.022500e+06	( <a href="#">53.8289048</a> , <a href="#">-1.8310423</a> )	<a href="#">53.828905</a>	-1.831042
196	ALBION SQUARE	2.450000e+06	(-41.27375755, 173.289393239104)	<a href="#">-41.273758</a>	<a href="#">173.289393</a>
197	ALBION STREET	2.096667e+06	( <a href="#">36.1659927</a> , -86.8074413)	<a href="#">36.165993</a>	-86.807441
391	ANHALT ROAD	2.435000e+06	( <a href="#">51.4803265</a> , -0.1667607)	<a href="#">51.480326</a>	-0.166761
406	ANSDELL TERRACE	2.250000e+06	( <a href="#">51.4998899</a> , -0.1891027)	<a href="#">51.499890</a>	-0.189103
421	APPLEGARTH ROAD	2.400000e+06	( <a href="#">53.749244</a> , -0.32678)	<a href="#">53.749244</a>	-0.326780
552	ASHCHURCH PARK VILLAS	2.150000e+06	( <a href="#">51.5000507</a> , -0.2421733)	<a href="#">51.500051</a>	-0.242173
671	AVENUE ROAD	2.143471e+06	( <a href="#">51.4067969</a> , -0.049519)	<a href="#">51.406797</a>	-0.049519
699	AYLESTONE AVENUE	2.286667e+06	( <a href="#">51.5409157</a> , -0.2178742)	<a href="#">51.540916</a>	-0.217874
760	BALLINGDON ROAD	2.105000e+06	( <a href="#">51.4541892</a> , -0.1588555)	<a href="#">51.454189</a>	-0.158856
853	BARONSMEAD ROAD	2.375000e+06	( <a href="#">51.4773147</a> , -0.239457)	<a href="#">51.477315</a>	-0.239457
979	BEAUCLERC ROAD	2.480000e+06	( <a href="#">51.4995771</a> , -0.2290331)	<a href="#">51.499577</a>	-0.229033
1089	BELSIZE CRESCENT	2.000000e+06	( <a href="#">51.5495675</a> , -0.1727325)	<a href="#">51.549568</a>	-0.172733
1100	BELVEDERE DRIVE	2.340000e+06	( <a href="#">51.4249173</a> , -0.2120774)	<a href="#">51.424917</a>	-0.212077
1132	BERESFORD TERRACE	2.100000e+06	( <a href="#">55.4571954</a> , <a href="#">-4.6291344</a> )	<a href="#">55.457195</a>	-4.629134
1190	BETTRIDGE ROAD	2.025000e+06	( <a href="#">51.471074</a> , -0.2028354)	<a href="#">51.471074</a>	-0.202835
1213	BICKENHALL STREET	2.208500e+06	( <a href="#">51.5211969</a> , -0.1589341)	<a href="#">51.521197</a>	-0.158934
1251	BIRCHLANDS AVENUE	2.217000e+06	( <a href="#">51.4483941</a> , -0.1604676)	<a href="#">51.448394</a>	-0.160468
1380	BLYTHS WHARF	2.000000e+06	(51.50900105, -0.0345837265300945)	51.509001	-0.034584
1431	BOSTON PLACE	2.167500e+06	(51.5073726, -0.340337807312508)	51.507373	-0.340338
1511	BRACKENBURY GARDENS	2.150000e+06	(51.5006233, -0.230729)	51.500623	-0.230729
1550	BRAMPTON GROVE	2.456875e+06	(51.5703648, -0.2833944)	51.570365	-0.283394
1555	BRAMSHOT AVENUE	2.177900e+06	(51.4811602, 0.0226516)	51.481160	0.022652
1629	BRANDAL F GARDENS	2.397132e+06	(51.5601748, -0.1954305)	51.560175	-0.195431

## Exploratory Data Analysis

Started analyzing the outcome of consolidated data to understand and observe the number of venues nearby streets. We need to see the relation between number of venues compare to house prices. We have around 9289 records associated to multiple venues.

Table 2. Venues Counts per Street

	Street Latitude	Street Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Street						
AIREDALE AVENUE	3	3	3	3	3	3
ALBION SQUARE	27	27	27	27	27	27
ALBION STREET	11	11	11	11	11	11
ANHALT ROAD	14	14	14	14	14	14
ANSDELL TERRACE	58	58	58	58	58	58
APPLEGARTH ROAD	4	4	4	4	4	4
ASHCHURCH PARK VILLAS	27	27	27	27	27	27
AVENUE ROAD	4	4	4	4	4	4
AYLESTONE AVENUE	4	4	4	4	4	4
BALLINGDON ROAD	14	14	14	14	14	14
BARONSMEAD ROAD	11	11	11	11	11	11
BEAUCLERC ROAD	29	29	29	29	29	29
BELSIZE CRESCENT	12	12	12	12	12	12
BELVEDERE DRIVE	38	38	38	38	38	38
BERESFORD TERRACE	26	26	26	26	26	26
BETTRIDGE ROAD	40	40	40	40	40	40
BICKENHALL STREET	100	100	100	100	100	100
BIRCHLANDS AVENUE	11	11	11	11	11	11
BLYTHS WHARF	16	16	16	16	16	16
BOSTON PLACE	18	18	18	18	18	18
BRACKENBURY GARDENS	16	16	16	16	16	16
BRAMPTON GROVE	4	4	4	4	4	4
BRAMSHOT AVENUE	7	7	7	7	7	7
BRIARDALE GARDENS	7	7	7	7	7	7
BRIDLANDS ROAD	2	2	2	2	2	2



Around 252 potential streets allocated. For better insight, we have calculated the average of venues per street.

Table 3. Average Venue

	Street	Accessories Store	Adult Boutique	Afghan Restaurant	African Restaurant	American Restaurant	Amphitheater	Antique Shop	Arcade
1	AIREDALE AVENUE	0.000000	0.0	0.0	0.000000	0.000000	0.000000	0.000000	0.0
2	ALBION SQUARE	0.000000	0.0	0.0	0.000000	0.000000	0.000000	0.000000	0.0
3	ALBION STREET	0.000000	0.0	0.0	0.000000	0.090909	0.090909	0.000000	0.0
4	ANHALT ROAD	0.000000	0.0	0.0	0.000000	0.000000	0.000000	0.000000	0.0
5	ANSDELL TERRACE	0.000000	0.0	0.0	0.000000	0.000000	0.000000	0.000000	0.0
6	APPLEGARTH ROAD	0.000000	0.0	0.0	0.000000	0.000000	0.000000	0.000000	0.0
7	ASHCHURCH PARK VILLAS	0.000000	0.0	0.0	0.000000	0.000000	0.000000	0.000000	0.0
8	AVENUE ROAD	0.000000	0.0	0.0	0.000000	0.000000	0.000000	0.000000	0.0
9	AYLESTONE AVENUE	0.000000	0.0	0.0	0.000000	0.000000	0.000000	0.000000	0.0

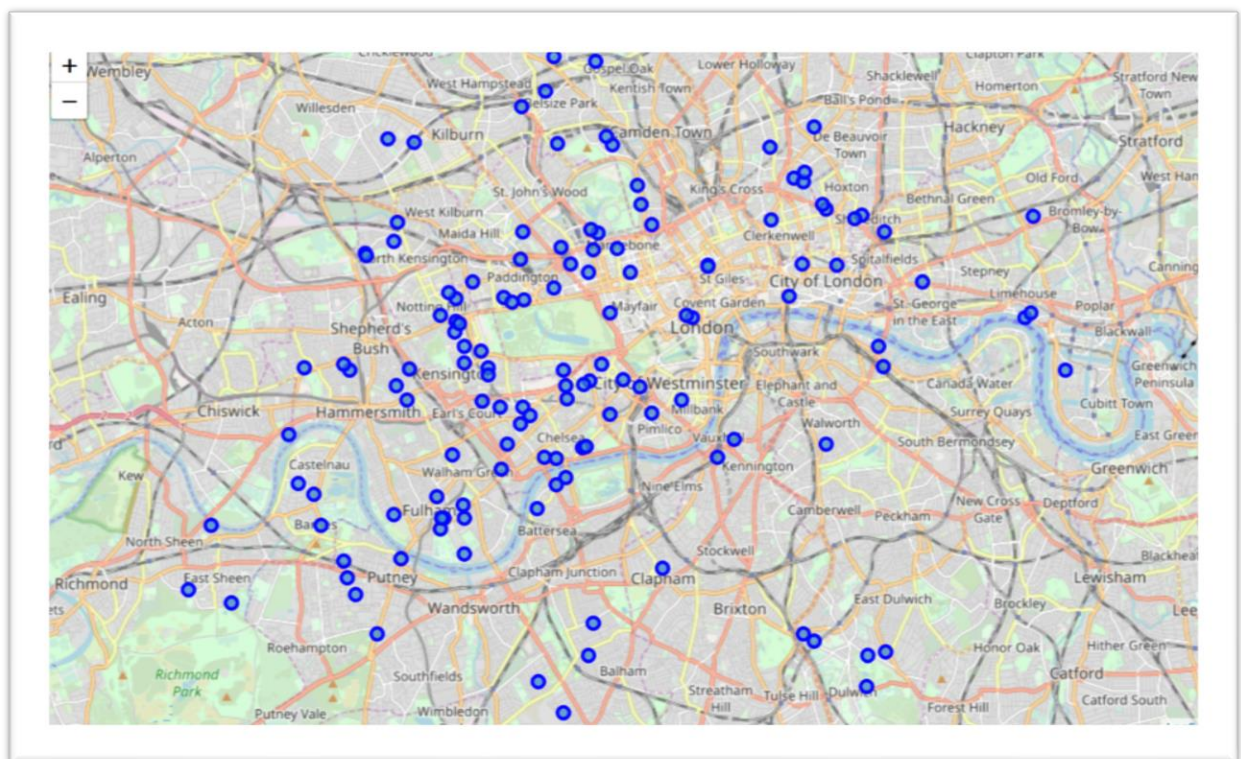
The most common top venues/facilities nearby profitable real estate investments are allocated for better visibility of homebuyer's interest.

Table 4. Top Venues

Street	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
AIREDALE AVENUE	Bakery	Harbor / Marina	Business Service	Zoo Exhibit	Food & Drink Shop	Fast Food Restaurant	Filipino Restaurant	Fish & Chips Shop	Fish Market
ALBION SQUARE	Café	Restaurant	Bar	Indian Restaurant	Coffee Shop	Pub	French Restaurant	Department Store	Supermarke
ALBION STREET	BBQ Joint	Pizza Place	Coffee Shop	Auto Workshop	Athletics & Sports	Lounge	Smoothie Shop	Amphitheater	American Restaurant
ANHALT ROAD	Pub	Grocery Store	Japanese Restaurant	Garden	Gym / Fitness Center	English Restaurant	Diner	Pizza Place	Cocktail Bar
ANSDELL TERRACE	Clothing Store	Italian Restaurant	Restaurant	Hotel	Juice Bar	Pub	Chinese Restaurant	English Restaurant	Indian Restaurant

Map visulaization is set to see the distribution of locations with respect to venues

Map 1. Street/ Venues

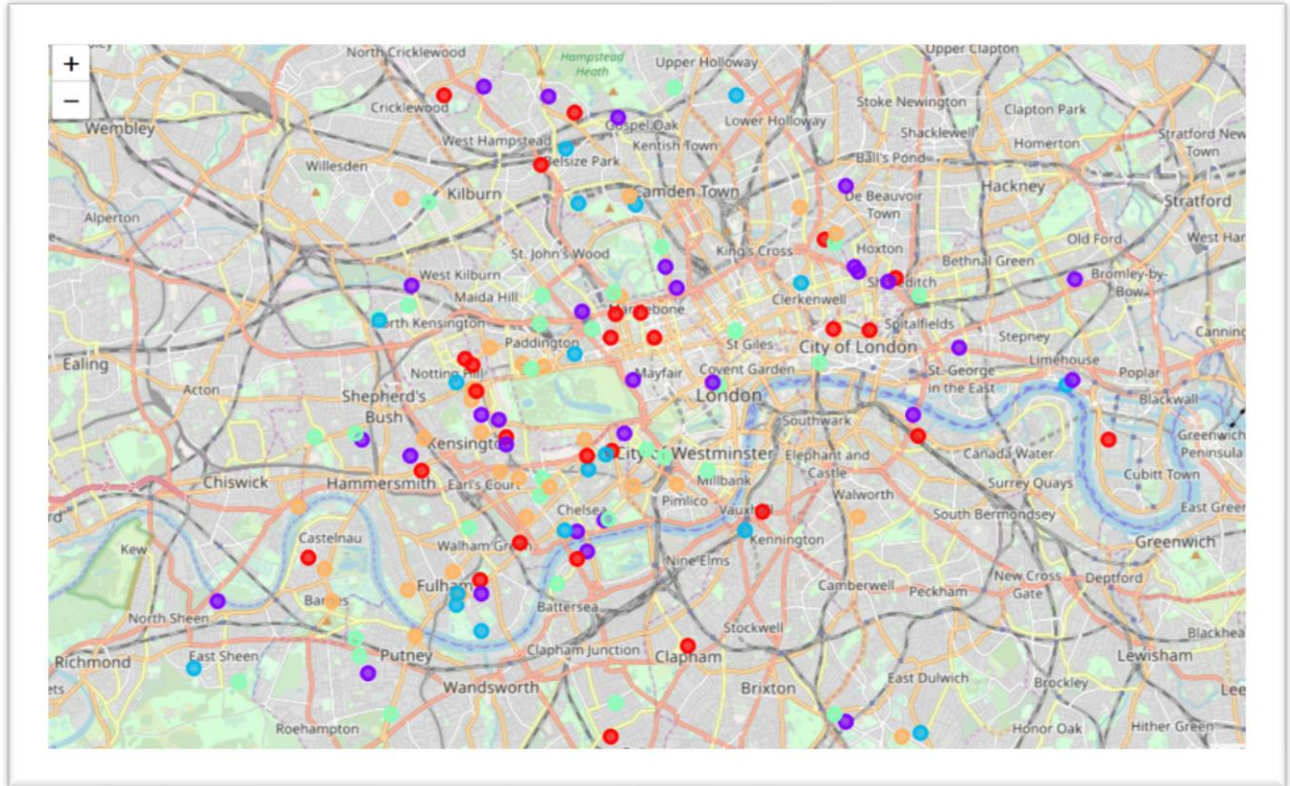




## 4. Predictive Modeling

The K-Mean Clustring model is being used for London Real Estate project. We have applied clustering method to help recommend best invistment and recommendation according to surrounding facilities. k-means clustering technique will be used to analyze real etates data for London. Before, venues around areas will be identified through foursequare API.

## Map 2. Clusters



With a population of more than 8.6 million, London is a densely populated metropolis with a melting pot of multi-ethnic residents from all over the world. As the hub for the UK's economy, politics and culture this city attracts a great deal of Brits and foreigners, despite high costs of living and higher than average housing prices than the rest of the UK.

Buying UK property is likely to be one of the biggest investments of your life, particularly in London where house prices are considerably higher. And with so many people all looking for accommodation in London, the housing market is very competitive for buying London property, with the market operating under its own influences and factors.

prices vary depending on where and the type of London property you buy.

In our analysis, outcome divided into 5 clustered included prices and common venues., Each cluster shows the most common venues targeted as example Clusters 4, indicated that target

home buyers live in 'green' areas with parks, theaters while cluster 0, target restaurants and cafes. The areas within Central London are often a preferred choice a close proximity to everything.

### Cluster No 0:

	Avg_Price	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
406	2.250000e+06	Clothing Store	Italian Restaurant	Restaurant	Hotel	Juice Bar	Pub	Chinese Restaurant	English Restaurant	Indian Restaurant	Garden
1213	2.208500e+06	Hotel	Café	Restaurant	Coffee Shop	Pizza Place	Gastropub	Pub	Chinese Restaurant	Indian Restaurant	Bakery
1251	2.217000e+06	French Restaurant	Pub	Lake	Train Station	Chinese Restaurant	Bakery	Coffee Shop	Breakfast Spot	Brewery	Filipino Restaurant
1863	2.197583e+06	Hotel	Middle Eastern Restaurant	Coffee Shop	Restaurant	Italian Restaurant	Sandwich Place	Pub	Café	Lebanese Restaurant	Chinese Restaurant
2148	2.188333e+06	Coffee Shop	Café	Italian Restaurant	Pizza Place	Grocery Store	Japanese Restaurant	Chinese Restaurant	Asian Restaurant	Sandwich Place	Hotel

### Cluster No 1:

	Avg_Price	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
196	2450000.0	Café	Restaurant	Bar	Indian Restaurant	Coffee Shop	Pub	French Restaurant	Department Store	Supermarket	Fish & Chips Shop
391	2435000.0	Pub	Grocery Store	Japanese Restaurant	Garden	Gym / Fitness Center	English Restaurant	Diner	Pizza Place	Cocktail Bar	Plaza
421	2400000.0	Pub	Nightclub	Casino	Zoo Exhibit	Food Court	Fast Food Restaurant	Filipino Restaurant	Fish & Chips Shop	Fish Market	Flea Market
979	2480000.0	Pub	Coffee Shop	Hotel	Thai Restaurant	Grocery Store	Ice Cream Shop	Chinese Restaurant	Fish & Chips Shop	Fish Market	Cocktail Bar
1550	2456875.0	Bar	Lake	Middle Eastern Restaurant	Men's Store	Friterie	Flower Shop	Farm	Farmers Market	Furniture / Home Store	Fast Food Restaurant

### Cluster No 2:

	Avg_Price	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
146	2022500.0	Bakery	Harbor / Marina	Business Service	Zoo Exhibit	Food & Drink Shop	Fast Food Restaurant	Filipino Restaurant	Fish & Chips Shop	Fish Market	Flea Market
1089	2000000.0	Italian Restaurant	History Museum	Movie Theater	Hotel	Bed & Breakfast	Café	Pub	Plaza	Greek Restaurant	Bakery
1190	2025000.0	Italian Restaurant	Café	Coffee Shop	Pub	Grocery Store	Park	French Restaurant	Yoga Studio	Bakery	Deli / Bodega
1380	2000000.0	Italian Restaurant	Pub	Convenience Store	Beer Garden	Gym	Canal Lock	Pizza Place	Athletics & Sports	Plaza	Turkish Restaurant
1700	2043000.0	Home Service	Supermarket	Zoo Exhibit	Food	Farm	Farmers Market	Fast Food Restaurant	Filipino Restaurant	Fish & Chips Shop	Fish Market



### Cluster No 3:

	Avg_Price	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
197	2.096667e+06	BBQ Joint	Pizza Place	Coffee Shop	Auto Workshop	Athletics & Sports	Lounge	Smoothie Shop	Amphitheater	American Restaurant	Shopping Mall
552	2.150000e+06	Grocery Store	Pub	Mediterranean Restaurant	Bakery	Indian Restaurant	Park	Coffee Shop	Café	Middle Eastern Restaurant	Wine Shop
671	2.143471e+06	Park	Tram Station	Grocery Store	Tapas Restaurant	Zoo Exhibit	Flea Market	Farm	Farmers Market	Fast Food Restaurant	Filipino Restaurant
760	2.105000e+06	Pub	Café	Accessories Store	Coffee Shop	Grocery Store	Bakery	Antique Shop	Sporting Goods Shop	Italian Restaurant	Women's Store
1132	2.100000e+06	Clothing Store	Bar	Platform	Coffee Shop	Italian Restaurant	Hotel	Music Store	Movie Theater	Supermarket	Shopping Mall

### Cluster No 4:

	Avg_Price	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
699	2.286667e+06	Park	Café	Movie Theater	Zoo Exhibit	Food Court	Fast Food Restaurant	Filipino Restaurant	Fish & Chips Shop	Fish Market	Flea Market
853	2.375000e+06	Movie Theater	Restaurant	Food & Drink Shop	Coffee Shop	Nature Preserve	Indie Movie Theater	Pub	Breakfast Spot	Thai Restaurant	Park
1100	2.340000e+06	Pub	French Restaurant	Coffee Shop	Bakery	Thai Restaurant	Italian Restaurant	Sushi Restaurant	Lounge	Mediterranean Restaurant	Scenic Lookout
2064	2.375000e+06	Pub	Park	Hotel	Italian Restaurant	Grocery Store	Indian Restaurant	Yoga Studio	Coffee Shop	Bakery	Bubble Tea Shop
2125	2.379653e+06	Pub	Bakery	Coffee Shop	Indian Restaurant	Grocery Store	Yoga Studio	Hotel	Park	Hostel	Record Shop

## 5. Conclusions

We have gathered data related to London properties including prices paid from Land Registry website. The places got explored across different locations in London and according to different venues including amenities and facilities. Data has been extracted from FourSquare APIs and got sorted out and arranged for visualization. Accordingly, we were able to recommend proper profitable real estate investment.

As a result, It has been found that areas like Notting Hill, Kensington, Marylebone, Brompton considered highly profitable venues to purchase a real estate according to amenities and essential facilities surrounding such venues i.e. Parks, supermarket, schools, hospitals, etc. In the other hand, Chelsea, Wandsworth, Balham and Fulham are next future elite venues with a wide range of amenities and facilities.

The result was divided into 5 clustered and it has been analyzed, each cluster shows the most common venues targeted as example Clusters 4, indicated that target home buyers live in 'green' areas with parks, theaters while cluster 0, target restaurants and cafes.

