

The Battle of Neighbourhoods

Background

According to Bloomberg News, the London Housing Market is in a rut. It is now facing several different headwinds, including the prospect of higher taxes and a warning from the Bank of England that U.K. home values could fall as much as 30 percent in the event of a disorderly exit from the European Union. More specifically, four overlooked cracks suggest that the London market may be in worse shape than many realize: hidden price falls, record-low sales, homebuilder exodus and tax hikes addressing overseas buyers of homes in England and Wales.

Business Problem

In this scenario, it is urgent to adopt machine learning tools in order to assist homebuyer's clientele in London to make wise and effective decisions. As a result, the business problem we are currently posing is: how could we provide support to homebuyer's clientele in 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 neighbourhoods 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 i.e. elementary schools, high schools, hospitals & grocery stores.

Data section

Data on London properties and the relative price paid data were extracted from the HM Land Registry (<http://landregistry.data.gov.uk/>). The following fields comprise the address data included in Price Paid Data: Postcode; PAON Primary Addressable Object Name. Typically, the house number or name; SAON Secondary Addressable Object Name. If there is a sub-building, for example, the building is divided into flats, there will be a SAON; Street; Locality; Town/City; District; County.

To explore and target recommended locations across different venues according to the presence of amenities and essential facilities, we will access data through FourSquare API interface and arrange them as a dataframe for visualization. By merging data on London properties and the relative price paid data from the HM Land Registry and data on amenities and essential facilities surrounding such properties from FourSquare API interface, we will be able to recommend profitable real estate investments.

Methodology section

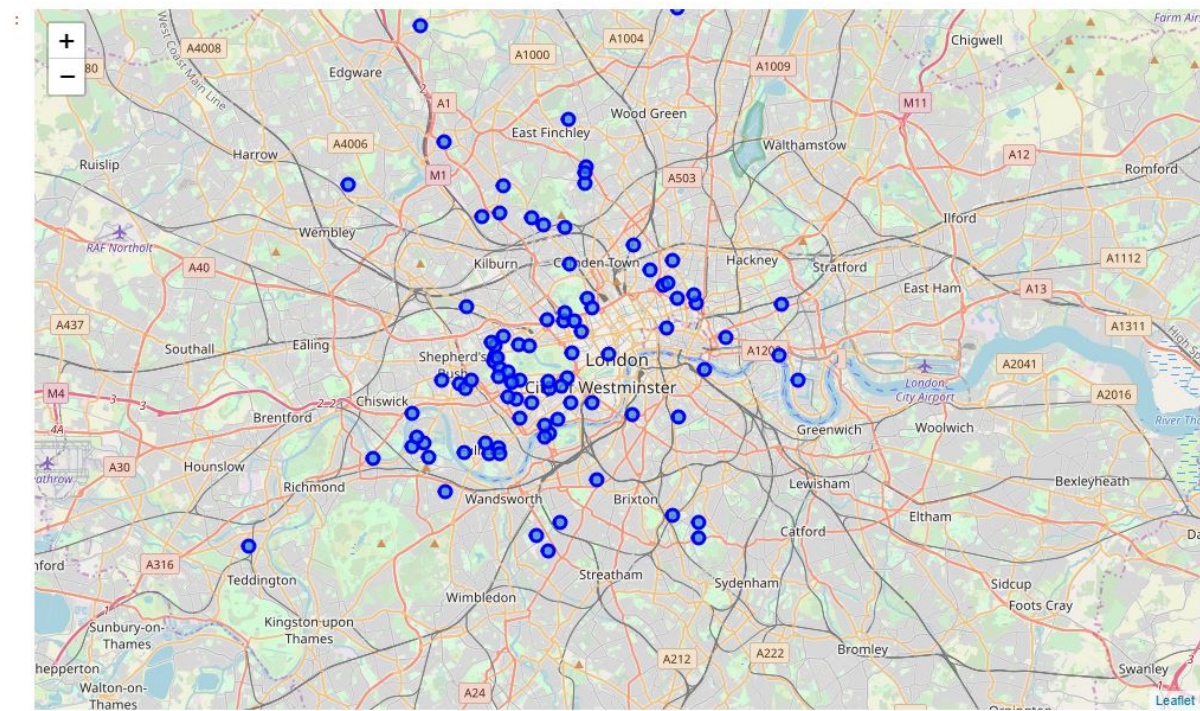
The Methodology section will describe the main components of our analysis and predication system. The Methodology section comprises four stages:

1. Collect Inspection Data
2. Explore and Understand Data
3. Data preparation and pre-processing
4. Modelling

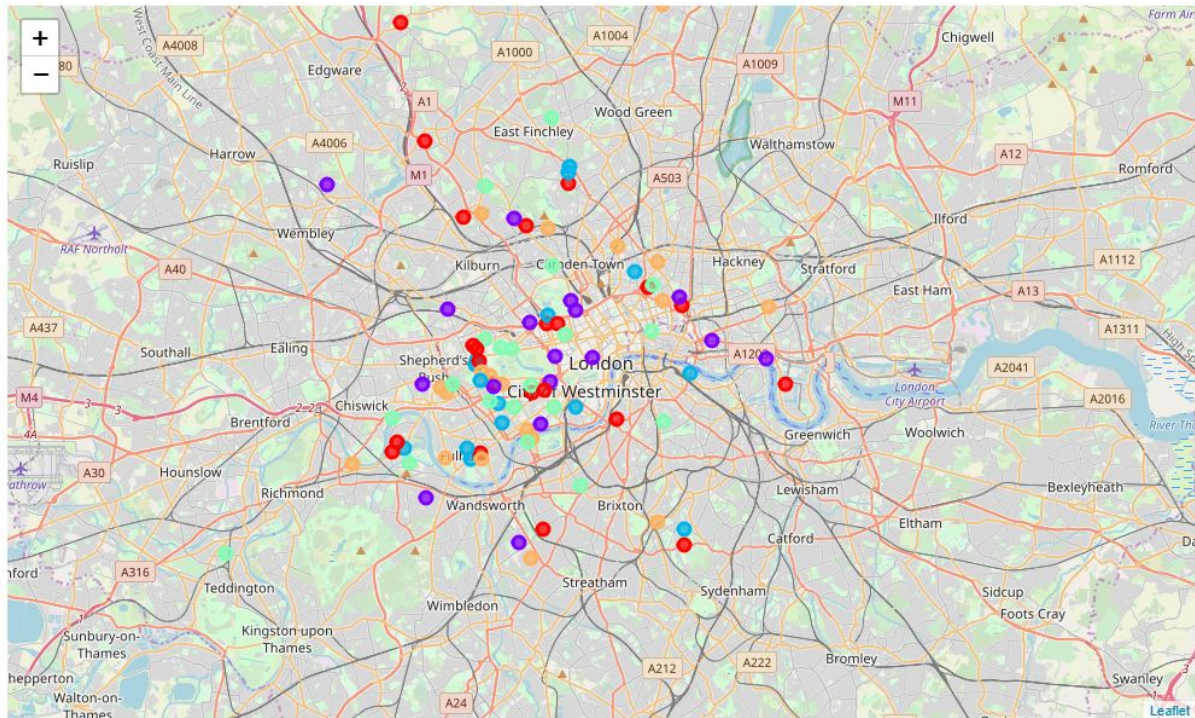
Prepared Data – Average By Price

	Street	Avg_Price	Latitude	Longitude
20	ABBOTSBURY CLOSE	2.367093e+06	51.532259	-0.006153
178	ALBION SQUARE	2.450000e+06	-41.273758	173.289393
355	ANHALT ROAD	2.435000e+06	51.480326	-0.166761
368	ANSDELL TERRACE	2.250000e+06	51.499890	-0.189103
381	APPLEGARTH ROAD	2.400000e+06	53.748654	-0.326670
617	AYLESTONE AVENUE	2.286667e+06	51.540916	-0.217874
753	BARONSMEAD ROAD	2.375000e+06	51.477315	-0.239457
867	BEAUCLERC ROAD	2.480000e+06	51.499577	-0.229033
1079	BICKENHALL STREET	2.351667e+06	51.521197	-0.158934
1094	BILLING ROAD	2.200000e+06	51.481883	-0.187862
1108	BIRCHLANDS AVENUE	2.217000e+06	51.448394	-0.160468
1310	BOWERDEAN STREET	2.300000e+06	51.472710	-0.192485
1371	BRAMPTON GROVE	2.475833e+06	51.570365	-0.283394
1439	BRIARDALE GARDENS	2.397132e+06	51.560175	-0.195431
1605	BROWNING CLOSE	2.320000e+06	51.885850	0.856081
1820	CALLCOTT STREET	2.375000e+06	51.508350	-0.198328
1871	CAMPDEN HILL ROAD	2.352889e+06	51.506461	-0.198896

Map of London



K-Means Clustering Map



Results – Discussion

First, even though the London Housing Market may be in a rut, it is still an "ever-green" for business affairs.

We may discuss our results under two main perspectives.

First, we may examine them according to neighbourhoods/London areas. It is interesting to note that, although West London (Notting Hill, Kensington, Chelsea, Marylebone) and North-West London (Hampstead) might be considered highly profitable venues to purchase a real estate according to amenities and essential facilities surrounding such venues i.e. elementary schools, high schools, hospitals & grocery stores, South-West London (Wandsworth, Balham) and North-West London (Islington) are arising as next future elite venues with a wide range of amenities and facilities. Accordingly, one might target under-priced real estates in these areas of London in order to make a business affair.

Second, we may analyse our results according to the five clusters we have produced. Even though, all clusters could praise an optimal range of facilities and amenities, we have found two main patterns. The first pattern we are referring to, i.e. Clusters 0, 2 and 4, may target home buyers prone to live in 'green' areas with parks, waterfronts. Instead, the second pattern we are referring to, i.e. Clusters 1 and 3, may target individuals who love pubs, theatres and soccer.

Conclusion

To sum up, according to Bloomberg News, the London Housing Market is in a rut. It is now facing several different headwinds, including the prospect of higher taxes and a warning from the Bank of England that U.K. home values could fall as much as 30 percent in the event of a disorderly exit from the European Union. In this scenario, it is urgent to adopt machine learning tools in order to assist homebuyer's clientele in London to make wise and effective decisions. As a result, the business problem we were posing was: how could we provide support to homebuyer's clientele in to purchase a suitable real estate in London in this uncertain economic and financial scenario?

To solve this business problem, we clustered London neighbourhoods in order to recommend venues and the current average price of real estate where homebuyers can make a real estate investment. We recommended profitable venues according to amenities and essential facilities surrounding such venues i.e. elementary schools, high schools, hospitals & grocery stores.

First, we gathered data on London properties and the relative price paid data were extracted from the HM Land Registry (<http://landregistry.data.gov.uk/>). Moreover, to explore and target recommended locations across different venues according to the presence of amenities and essential facilities, we accessed data through FourSquare API interface and arranged them as a data frame for visualization. By merging data on London properties and the relative price paid data from the HM Land Registry and data on amenities and essential facilities surrounding such properties from FourSquare API interface, we were able to recommend profitable real estate investments.

Second, The Methodology section comprised four stages: 1. Collect Inspection Data; 2. Explore and Understand Data; 3. Data preparation and pre-processing; 4. Modelling. In the modelling section, we used the k-means clustering technique as it is fast and efficient in terms of computational cost, is highly flexible to account for mutations in real estate market in London and is accurate.

Finally, we drew the conclusion that even though the London Housing Market may be in a rut, it is still an "ever-green" for business affairs. We discussed our results under two main perspectives. First, we examined them according to neighbourhoods/London areas. although West London (Notting Hill, Kensington, Chelsea, Marylebone) and North-West London (Hampstead) might be considered highly profitable venues to purchase a real estate according to amenities and essential facilities surrounding such venues i.e. elementary schools, high schools, hospitals & grocery stores, South-West London (Wandsworth, Balham) and North-West London (Islington) are arising as next future elite venues with a wide range of amenities and facilities. Accordingly, one might target under-priced real estates in these areas of London in order to make a business affair. Second, we analysed our results according to the five clusters we produced. While Clusters 0, 2 and 4 may target home buyers prone to live in 'green' areas with parks, waterfronts, Clusters 1 and 3 may target individuals who love pubs, theatres and soccer.