

Capstone Project Report - The Battle of Neighborhoods

Introduction

Background

According to Bloomberg News, the London Housing Market is in a rut. It is now facing a number of 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 homebuyers 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 homebuyers 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 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 i.e. elementary schools, high schools, hospitals & grocery stores.

Data

Data on London properties and the relative price paid data were extracted from the HM Land Registry (<http://landregistry.data.gov.uk/>).

Solution

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

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 preprocessing
4. Modeling

Data Processing

At this stage, we prepare our dataset for the modeling process, opting for the most suitable machine learning algorithm for our scope. Accordingly, we perform the following steps:

- Sort data by date of sale
- Select data only for the city of London
- Make a list of street names in London
- Calculate the street-wise average price of the property
- Read the street-wise coordinates into a data frame, eliminating recurring word London from individual names
- Join the data to find the coordinates of locations which fit into client's budget
- Plot recommended locations on London map along with current market prices

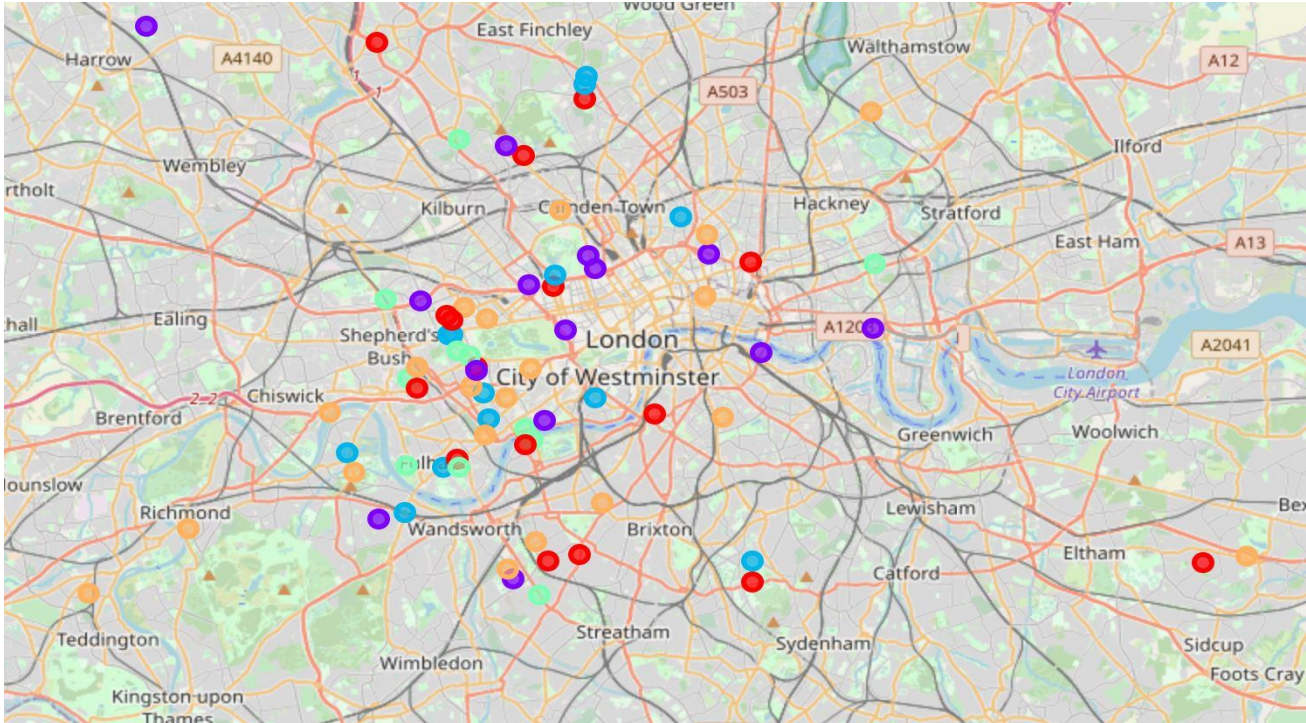
```
-----ABBOTSBURY CLOSE-----
      venue  freq
0  Thai Restaurant  0.2
1           Park    0.2
2   Waterfront    0.2
3  Grocery Store   0.2
4         Hotel    0.2
```

```
-----ALBION SQUARE-----
      venue  freq
0         Café  0.22
1          Bar  0.07
2          Pub  0.07
3 Indian Restaurant  0.07
4   Coffee Shop  0.07
```

```
-----ANHALT ROAD-----
      venue  freq
0          Pub  0.29
1 Japanese Restaurant  0.07
2 Gym / Fitness Center  0.07
3   Pizza Place  0.07
4         Diner  0.07
```

```
-----ANSDELL TERRACE-----
      venue  freq
0 Clothing Store  0.07
1 Italian Restaurant  0.07
2       Juice Bar  0.05
3         Hotel  0.05
4 Indian Restaurant  0.05
```

Clustering and Analysis



Ideal Localities

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
ABBOTSBURY CLOSE	Grocery Store	Park	Waterfront	Hotel	Thai Restaurant	Farm	Eastern European Restaurant	Electronics Store	English Restaurant
ALBION SQUARE	Café	Restaurant	Indian Restaurant	Bar	Coffee Shop	Pub	New American Restaurant	Seafood Restaurant	Fish & Chips Shop
ANHALT ROAD	Pub	Plaza	Pizza Place	Grocery Store	Japanese Restaurant	French Restaurant	English Restaurant	Gym / Fitness Center	Diner
ANSDELL TERRACE	Clothing Store	Italian Restaurant	Café	English Restaurant	Pub	Juice Bar	Hotel	Indian Restaurant	Bakery
APPLEGARTH ROAD	Bar	Pub	Casino	Nightclub	Fast Food Restaurant	English Restaurant	Event Space	Exhibit	Falafel Restaurant

Results and Discussion

First of all, 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 neighborhoods/London areas. It is interesting to note that, although West London (Notting Hill, Kensington, Chelsea, Marylebone) and North-West London (Hampsted) 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 analyze 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

Even though the London Housing Market may be in a bad shape, it is still an "ever-green" for business affairs.