<div align="center">

Omar Ragab
August 5, 2020

</div>

Peer-graded Assignment: Capstone Project - (The Battle of Neighborhoods (Week 1

Clearly define a problem or an idea of your choice, where you would need to leverage the Foursquare location data to solve or execute. Remember that data science problems always target an audience and are meant to help a group of stakeholders solves a problem, so make sure that you explicitly describe your audience and why they would .care about your problem

Introduction section

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

The 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 I am currently posing is: how could I provide support to homebuyers clientele in to purchase a ?suitable real estate in London in this uncertain economic and financial scenario

solving this problem

I am 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. I 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, I 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, I will be able to recommend profitable real estate investments.

Methodology section

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

Results and Discussion section

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

First, I will 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 (Isliington) 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, I will analyze our results according to the five clusters we have produced. Even though, all clusters could praise an optimal range of facilities and amenities, I have found two main patterns. The first pattern I am 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 I am 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 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. 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 I was posing: how could I 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, I clustered London neighborhoods in order to recommend venues and the current average price of real estate where homebuyers can make a real estate investment.I recommended profitable venues according to amenities and essential facilities surrounding such venues i.e. elementary schools, high schools, hospitals & grocery stores.

First, I 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, I 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 preprocessing; 4. Modeling. In particular, in the modeling 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, I drew the conclusion that even though the London Housing Market may be in a rut, it is still an "ever-green" for business affairs. I discussed our results under two main perspectives. First, I examined them according to neighborhoods/London areas. 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 (Isliington) 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, I analyzed our results according to the five clusters I 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.

```
dt_affordable
```

	Street	Avg_Price
196	ALBION SQUARE	2.450000e+06
390	ANHALT ROAD	2.435000e+06
405	ANSDELL TERRACE	2.250000e+06
422	APPLEGARTH ROAD	2.400000e+06
856	BARONSMEAD ROAD	2.375000e+06
982	BEAUCLERC ROAD	2.480000e+06
1103	BELVEDERE DRIVE	2.340000e+06
1216	BICKENHALL STREET	2.208500e+06
1254	BIRCHLANDS AVENUE	2.217000e+06
1554	BRAMPTON GROVE	2.456875e+06
1633	BRIARDALE GARDENS	2.397132e+06
1798	BROOKWAY	2.400000e+06
1915	BURBAGE ROAD	2.445000e+06
1981	BURY WALK	2.492500e+06
2069	CALLCOTT STREET	2.375000e+06
2130	CAMPDEN HILL ROAD	2.379653e+06
2137	CAMPION ROAD	2.461000e+06