**Introduction**

According to news agency, London’s housing market is in a rut. It’s facing lots of issues, with a possible increase in taxes and a warning from Bank of England that housing values will drop in the event of a disorderly exit from EU. There are some other issues that may caused this downfall, which includes hidden price fall, record-low sales, homebuilder exodus and hike in taxes for overseas buyers.

**Business Problem**

In this scenario, it is urgent to adopt machine learning tools in order to assist home buyers 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 home buyers 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**

Data on London properties and the relative price paid data were extracted from the HM Land Registry. 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 Four Square API interface and arrange 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 Four Square API interface, we will be able to recommend profitable real estate investments.

**Solution**

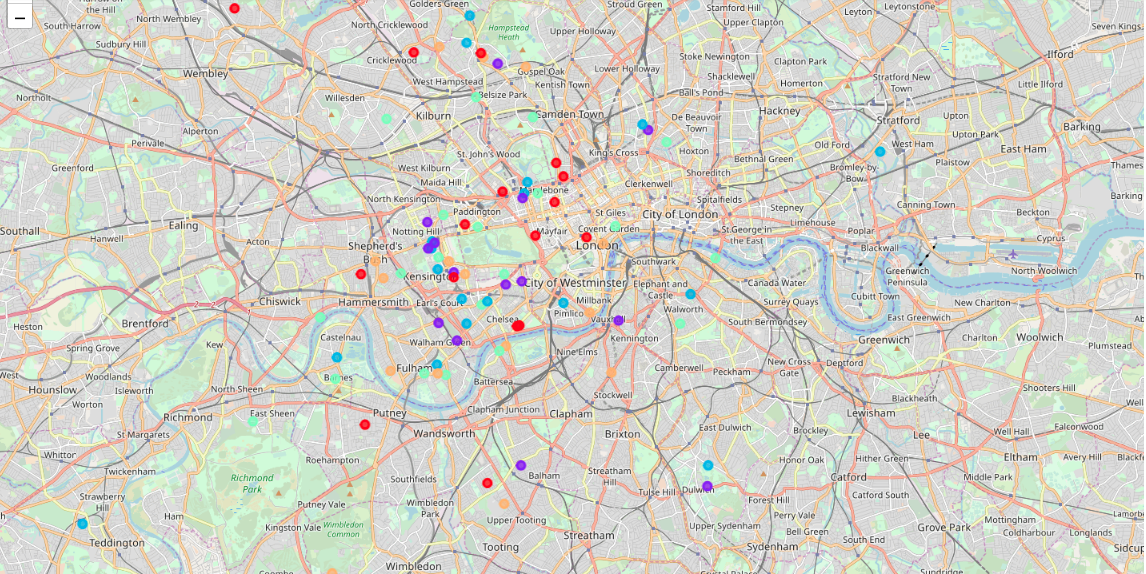
Clustering London neighbourhoods in order to recommend venues and the current average price of real estate where homebuyers can make a real estate investment

**Methodology**

This section comprises four different stages:

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

**K-Means Clustering**



**Outcome**

* West London (Notting Hill, Kensington, Chelsea, Marylebone) and North-West London (Hampstead) might be considered highly profitable venues to purchase a real estate;
* 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.
* Examination of real estates by clusters
* 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.