Background

According to reports, the London Housing Market is now facing a number of different headwinds. 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. Hidden price falls, record-low sales, homebuilder exodus and tax hikes addressing overseas buyers of homes in England and Wales may lead to falling house prices. There's a sense that political uncertainty continues to stifle growth.

Business Problem¶

In this scenario, machine learning tools are very critical in order to assist homebuyers clientele in London in making wise and effective decisions. As a result, the business problem we are currently posing is: How could we provide support to homebuyers in to purchase a suitable real estate in London in this uncertain economic and financial scenario?

To solve this business problem, I will be 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, 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
The Methodology section comprises four stages:

1. Collect Inspection Data
2. Explore and Understand Data
3. Data preparation and preprocessing
4. Modeling