Business Problem:

The city of London is one of the oldest cities in the world with one of the most cosmopolitan living environments. The population within the London commuter belt is second highest in the EU with 14,040,163 inhabitants in 2016. Close to 300K people move to London every year with an intention to stay short term or long term. The London Real Estate, catering to housing and commercial requirements has been facing many challenges recently such as new real estate tax rules and fall in property prices due to slow economic conditions resulting from uncertain business conditions caused by Brexit. Moreover London is wide spread with many options of new and old building projects each having its own pros and cons. How can we provide support to homebuyers clientele to purchase a suitable real estate in London in this uncertain economic and financial scenario?

Solution:

Clustering London neighbourhoods in terms of facilities and current average property prices to help homebuyers make an informed investment in the real estate of the city.

Data and Methodology:

Data:
- List of areas of London with its boroughs and postcodes.
  (https://en.wikipedia.org/wiki/List_of_areas_of_London)

- Data on London property prices from the HM Land Registry (
  http://landregistry.data.gov.uk/)

- Data on amenities and essential facilities surrounding properties from FourSquare API interface.

Mehodology:

- *Collect Inspection Data*: Downloading it from HM Land Registry website.

```
df_ppd.head(5)
```

ut[3]:

	{666758D7-43A9-3363-E053-6B04A8C0D74E}	405000	2018-01-25 00:00	WR15 8LH	D	N	F	RAMBLERS WAY	Unnamed: 8	Unnamed: 9	BORASTON	TENBURY WELLS	SHROPSHIRE	SHROPSHIRE.1	A
0	{666758D7-43AA-3363-E053-6B04A8C0D74E}	315000	2018-01-23 00:00	SY7 8QA	D	N	F	MONT CENISE	NaN	NaN	CLUN	CRAVEN ARMS	SHROPSHIRE	SHROPSHIRE	A
1	{666758D7-43AD-3363-E053-6B04A8C0D74E}	165000	2018-01-19 00:00	SY1 2BF	T	Y	F	42	NaN	PENSON WAY	NaN	SHREWSBURY	SHROPSHIRE	SHROPSHIRE	A
2	{666758D7-43B0-3363-E053-	370000	2018-01-22	SY8 4DF	D	N	F	WILLOW HEY	NaN	NaN	ASHFORD CARBONEL	LUDLOW	SHROPSHIRE	SHROPSHIRE	A

- *Explore and Understand Data*: Assigning meaningful row and column names, calculate average street price and map this data against coordinates.

df_affordable

9]:

	Street	Avg_Price
196	ALBION SQUARE	2.450000e+06
391	ANHALT ROAD	2.435000e+06
406	ANSDELL TERRACE	2.250000e+06
422	APPLEGARTH ROAD	2.400000e+06
855	BARONSMEAD ROAD	2.375000e+06
981	BEAUCLERC ROAD	2.480000e+06
1102	BELVEDERE DRIVE	2.340000e+06
1215	BICKENHALL STREET	2.208500e+06
1253	BIRCHLANDS AVENUE	2.217000e+06
1553	BRAMPTON GROVE	2.456875e+06
1632	BRIARDALE GARDENS	2.397132e+06

- *Data preparation and pre-processing:* Using Geocoder Library to find latitude and longitude of Columns. Add these Lats and Longs into Data Frame and create a map of London.

	Location	London_borough	Postcode district	Avg price	Latitude	Longitude
0	Abbey Wood	Bexley	SE2	368046	51.492450	0.121270
1	Crossness	Bexley	SE2	368046	51.492450	0.121270
2	West Heath	Bexley	SE2	368046	51.492450	0.121270
3	Acton	Ealing	W3	529790	51.513240	-0.267460
4	Addington	Croydon	CR0	363083	51.384755	-0.051499

> *Modelling K-Means Clustering*: Analyse neighbourhoods and cluster them as suitable venues according to amenities and essential facilities surrounding venues such as elementary schools, high schools, hospitals & grocery stores

```
----Alperton----
                 venue  freq
0                Hotel  0.07
1              Theater  0.05
2               Garden  0.04
3                Plaza  0.04
4  Monument / Landmark  0.03


----Anerley----
                 venue  freq
0           Supermarket  0.33
1     Convenience Store  0.17
2   Fast Food Restaurant  0.17
3                 Hotel  0.17
4         Grocery Store  0.17
```

Outcome:

Examination of Real Estates according to neighbourhoods/London areas:

- West London (Notting Hill, Kensington, Chelsea, Marylebone) and North-West London (Hampsted) might be considered highly profitable venues to purchase a real estate.
- 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. They are suitable for long term real estate investments.

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.

	Location	London_borough	Post town	Postcode district	Avg price	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue
11	Archway	Islington	LONDON	N19	572984	51.563930	-0.129450	5	Grocery Store	Coffee S
13	Arkley	Barnet	BARNET	EN5	630230	51.644415	-0.179185	5	Pub	Coffee S
15	Balham	Wandsworth	LONDON	SW12	665289	51.448220	-0.148390	5	Coffee Shop	Grocery Store
16	Bankside	Southwark	LONDON	SE1	923979	51.499600	-0.096130	5	Pub	Coffee S
24	Battersea	Wandsworth	LONDON	SW11	894664	51.467600	-0.162900	5	Café	Indian Restaur