**Context:**

The city of London was built in the year 43 AD by the Roman empire and is a city rich in history and geographical diversity. It has seen wars, plagues, fires and most recently a virus pandemic. As the capital city of the United Kingdom, it is home to 9 million residents and is a global hub for commerce and culture. Each year the city attracts tens of millions of visitors who come to marvel at historic sites such the Tower of London and Buckingham palace. The population of the city continues to grow each year as it attracts many migrants from around the world, in search of employment and a cosmopolitan lifestyle.

However, for all its charm the city has a fundamental problem. There is a severe shortage of housing in the city, and current supply is unable to keep up with the growth in demand.

The city is surrounded by two circular ring roads and a river which runs right through the heart of the city. Much of the infrastructure which supports the city was build centuries ago, and although it plays a key role in the modern world, the city stands on ancient foundations. To complicate this further, large parts of the city were either destroyed completely or badly bombed during the second world war.

Due to the high levels of investment, neighbourhoods are constantly being redeveloped and changed, new bars and restaurants open on a weekly basis, and glitzy penthouse apartments stand next to centuries old manor houses. This has created a challenging situation for those looking to enter the London property market, as affordable neighbourhoods are gentrified, and local residents are priced out.

**Problem:**

Due to the international investment and the high levels of demand for housing, the challenge faced by many of London’s inhabitants is the inability to afford to live in the very city they work in. When choosing a location to live, various factors come in to play depending on the individual. Some may be looking for a neighbourhood with lots of bars and restaurants, and a busy nightlife scene. Others may be looking for parks and outdoor spaces.

In this analysis I will cluster London neighbourhoods based their local features. The feature dataset will be obtained using the foursquare API service to extract a list of the venues nearby. This dataset will be used to group areas by their geographical similarity. I will also bring in data on house price transactions in the local area to understand local house prices and compare this for the different neighbourhoods.

The problem I aim to solve in this analysis is to identify the more affordable neighbourhoods that offer the same local venues and services as others. This will help anyone who lives in the city, or is looking to move to the city, find an affordable neighbourhood. Given the rapid rate of change in London, this will help those living in the city keep up with the changes and find affordable accommodation without compromising on their standard of living.

**Data:**

**London geospatial data** – In order to correctly identify the various different Geographical layers in London, I will require data on the latitudes and longitudes of each polygon. This dataset will be extracted from a .csv file available on the website <https://www.doogal.co.uk/london_postcodes.php>

**Foursquare data** – The foursquare api will be used to retrieve data on the characteristics of each neighbourhood. The data will retrieved using specific latitudes and longitudes corresponding to each area. This data will form the features dataset that will be used to cluster each neighbourhood.

**Transaction data from HM land registry** – In order to understand local house prices I will need data on UK house transactions. I will obtain this from the official government records at HM land registry found at: <http://prod.publicdata.landregistry.gov.uk.s3-website-eu-west-1.amazonaws.com/pp-2020.csv>. This data set includes a list of transactions across the United Kingdom since the start of 2020.