IBM Capstone Project: The Battle of Neighbourhoods

**Analysis of London Boroughs for Real Estate Investment**

Daniele Carli

2019-2020

**Introduction/Business problem**

A prominent UK holding company has decided to invest part of its capital into the real estate business. More specifically, the company is looking for buying multiple houses/apartment buildings in London.

Thus, the problem consists now in identifying which areas of the city are more suitable and advantageous for this kind of investment, which means understanding which areas have the optimal and more profitable characteristics for long-term or short-term house renting and, possibly, selling at a higher cost to potential buyers in the future.

**1. Data acquisition**

Three determinants have been identified as particularly influential in the choice of the locations where to acquire real estate properties:

- Level of criminality: a lower level of criminality could be determinant in the decision by a potential tenant/buyer to permanently live in a certain area. Therefore, it might constitute also a determinant for the willingness by those people to pay a higher price to live in a location where they can feel safe.

- Relevant nearby venues: once the borough with low crime level have been found, it is important to identify the most relevant venues in the surroundings. In fact, as easily perceivable, potential buyers or tenants would be more prone to live in an area characterized by high standard shops, services or recreational venues.

For the first determinant, I will retrieve information from the *MPS Borough Level Crime* dataset provided by the Metropolitan Police Service for the period between 1/07/2018 and 30/06/2020. The Database is available at this web page: <https://data.london.gov.uk/dataset/recorded_crime_summary>

With regard to the venues and related information about them I will resort to the Fourthsquare API. The objective is to explore the main facilities and services around the boroughs with lower crime rate. Subsequently, I will perform a quality check by analysing their average rate.

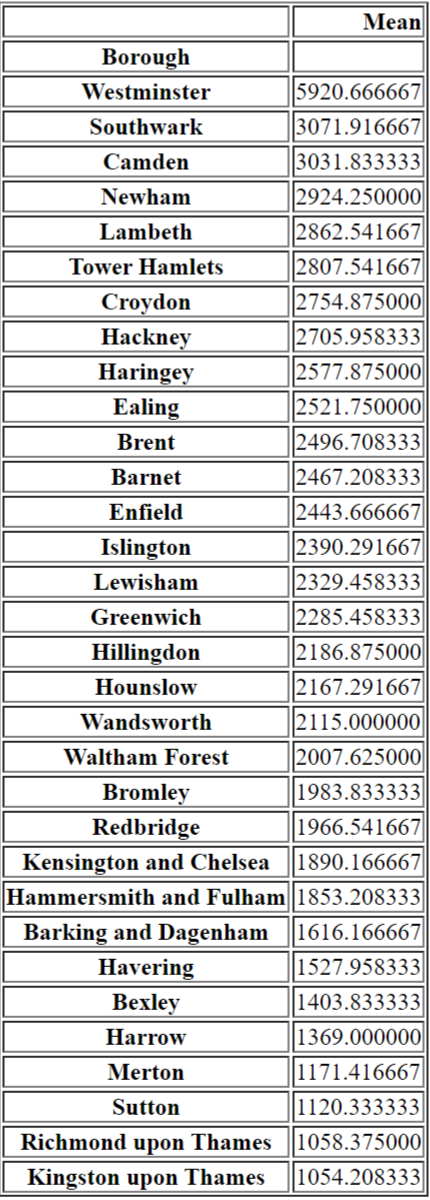
Ultimately, the project should provide to the client the top 3 London Boroughs for profitable real estate investment and some useful nearby tips in order to choose the optimal location into the boroughs

**2. Exploratory Data Analysis**

**2.1 Top 3 Boroughs for low crime rate**

The *MPS Borough Level Crime* dataset is originally organized as a register of each crime committed in London over 24 months (from 1/07/2018 and 30/06/2020). Since our purpose is to capture the crimes committed in each borough regardless of the typology of the crime, I wrangle the dataset in order to serve this purpose. Thus, crimes get grouped by borough throughout the *groupby* function and, then, a column called mean gets created, which is characterized by the mean of each borough crimes over 24 months. Finally, I use the *sort.values* function in order to sort the mean values in descending order.

Table 1: London Boroughs Average Crime per month



For the sake of clarity, a horizontal bar chart and a choropleth map of London might be useful in order to better corroborate the results of the analysis.

Figure 1: bar chart of London Boroughs monthly average crime rate

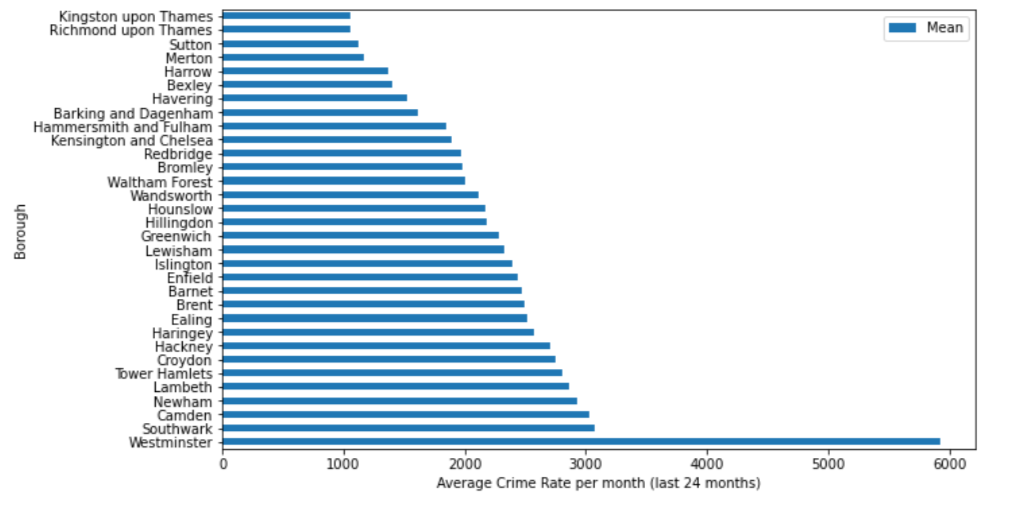
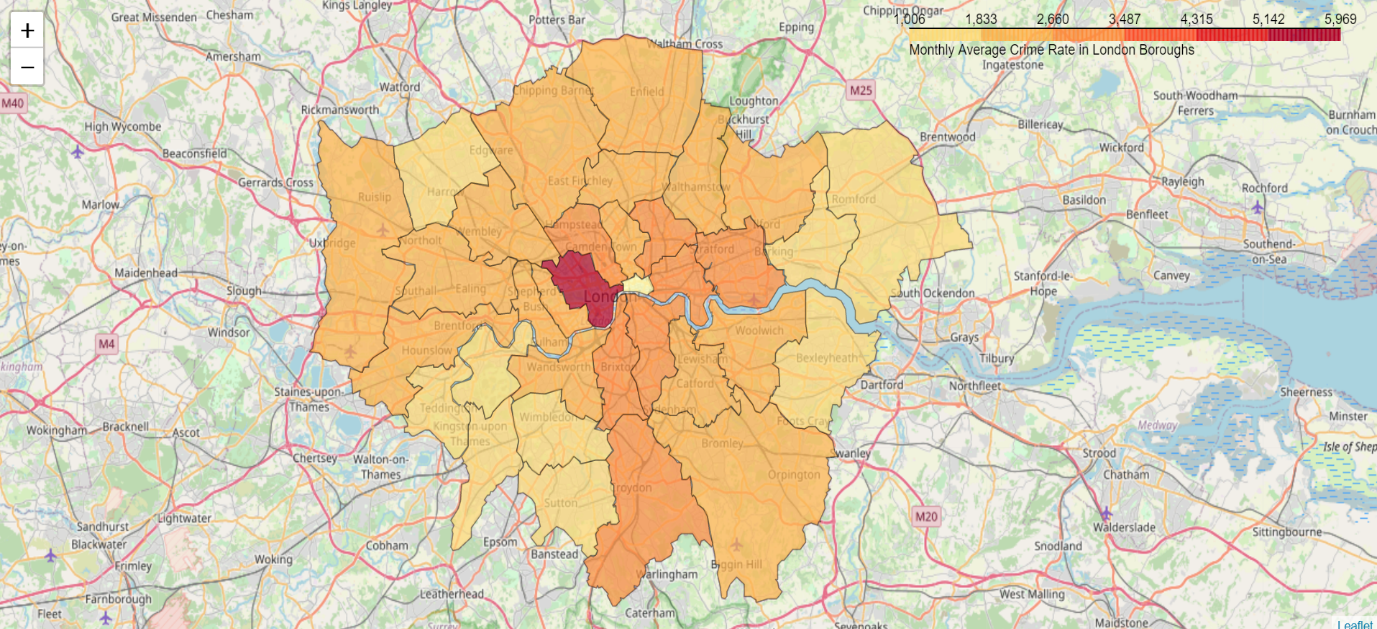


Figure 2: choropleth map of London on a borough-level



Two aspects easily emerge from the graphic representations below:

* *Westminster*, located in the centre of London, with almost 6000 crimes committed every month is the most unsafe borough in the UK capital, which is understandable considering the density of people and services that converge there every day.
* *Kingston Upon Thames*, *Richmond Upon Thames* and *Sutton* are characterized by the lowest monthly crime rates in London. For this reason, they represent an excellent top 3 boroughs for safety and, in turn, the best boroughs where to invest in real estate properties.

Having said that, I focus on these 3 boroughs for the last part of this analysis.

**2.2 Relevant nearby venues in the top 3 Boroughs**

Now that three optimal boroughs for safety have been identified, I want to identify which venues are more relevant in each borough and where they are located. For doing this, I resort to the Fourthsquare API, in particular we send multiple *Explore* request to the location data server which will return us two venues for each of the following big Fourthsquare venue categories that I have chosen:

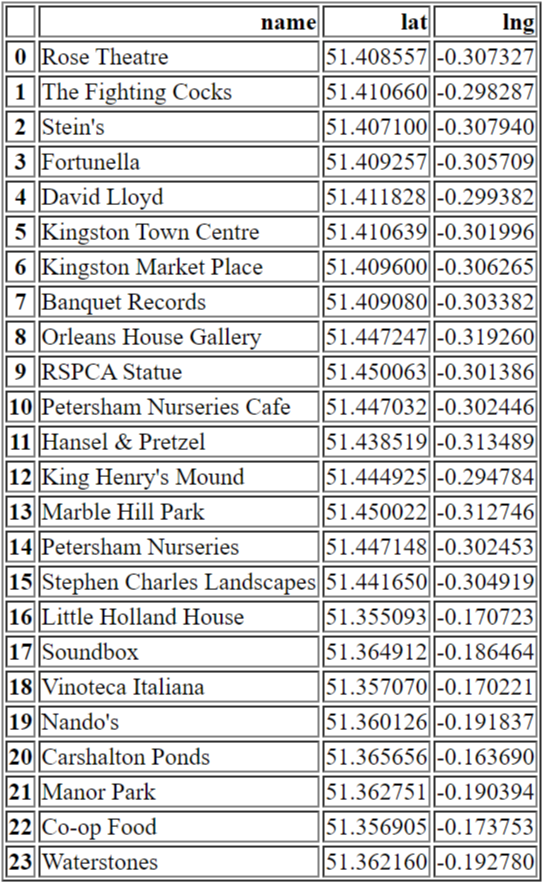
*- Arts and Entertainment*

*- Food*

*- Outdoor and Recreation*

*- Shops and Services*

Table 2: Relevant nearby Venues in top 3 London Boroughs



Ultimately, I create a *Folium* map with markers for each venue location. Here is the map divided for the three boroughs:

Figure 3: relevant nearby venues in Kingston Upon the Thames

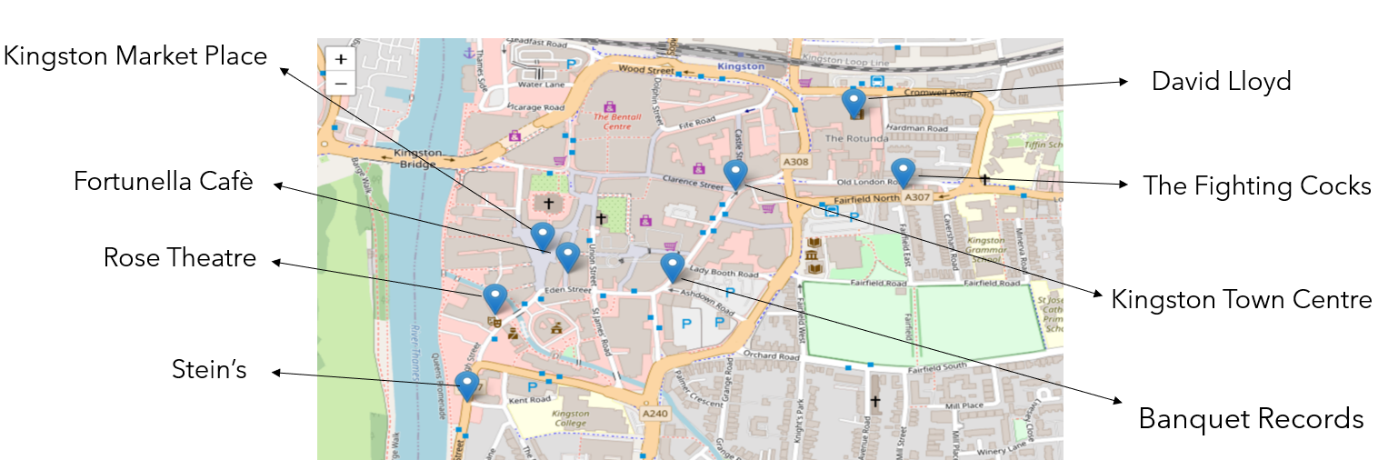
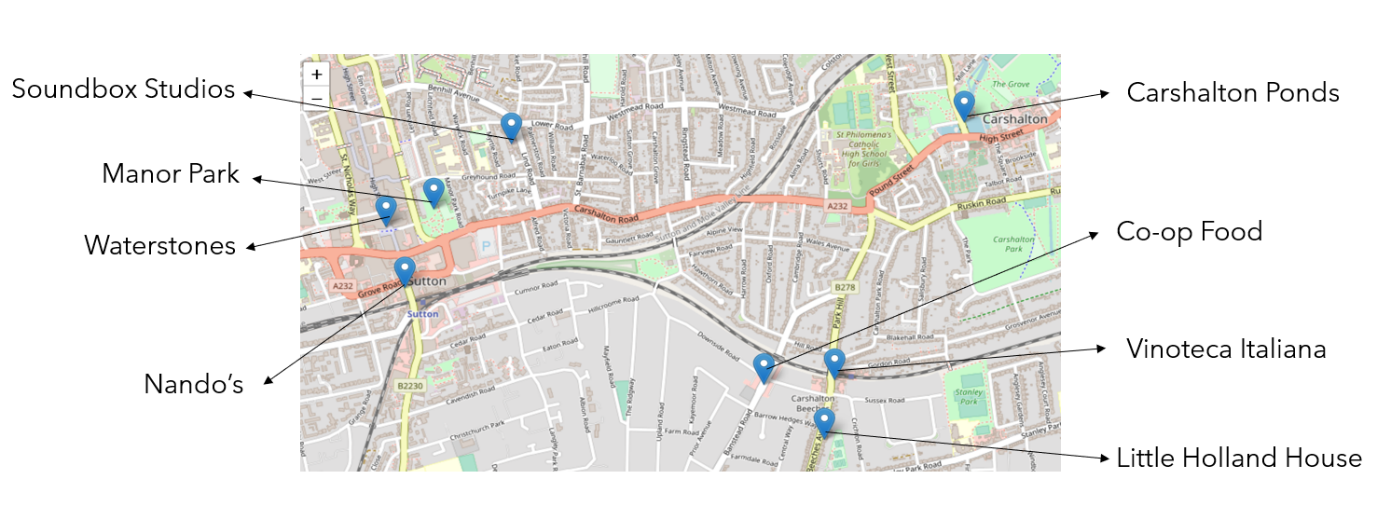


Figure 4: levant nearby venues in Richmond Upon the Thames 

Figure 5: relevant nearby venues in Sutton



**Conclusion**

This project had the purpose of providing to the client a convincing and data-driven analysis of the best areas in London where to invest in real estate properties. Since the beginning, I tried to identify myself with the potential future buyer or tenant, and from this I identified two fundamental determinants in the decision about where to reside: safety and strategic proximity to some venues of interest.

For the first variable I chose the monthly average crime rate in each London borough as an accurate indicator. The results of the analysis have shown that three boroughs can be considered the safest of the city, namely Kingston Upon Thames, Richmond Upon Thames and Sutton.

Then, in order to satisfy the second determinant of choice, I thought it could be useful for the client to get an overview of which venues thereabouts each of the top 3 Boroughs are more relevant; in fact, these information would definitely give a more accurate suggestion on which specific areas might be more desirable by buyers/tenants and, thus, more profitable for the owner. On this purpose, I resorted to the Fourthsquare API which allows to access this kind of location data through the *explore* function. As previously shown in the maps, 8 recommended venues for each of the Top 3 Boroughs have been selected by this analysis, including places such as Marble Hill Park, Kingston Market Place, Vinoteca Italiana and Little Holland House.

The hope is that the analytic process of this project, together with the support of graphic content, will guide the client to the optimal investment.