**Painty Studios LTD: Choosing a Location for an Art Gallery in London**

Alicia Davies

November 2020

**Introduction**

Painty Studios Ltd are looking to expand and set-up their first art gallery in London. The company is currently based in Finland and has little experience in the UK. They need to understand the differences between the London Boroughs and where the local population might have most interest in an Art Gallery. The more interest there is in the Art Gallery the more visitors they will expect to have, and the more art they can sell. Therefore, they have tasked their chief data scientist to analyse the different London Boroughs to suggest the most suitable location for their art gallery. The analysis from this study will be provided to the CEO for the final decision.

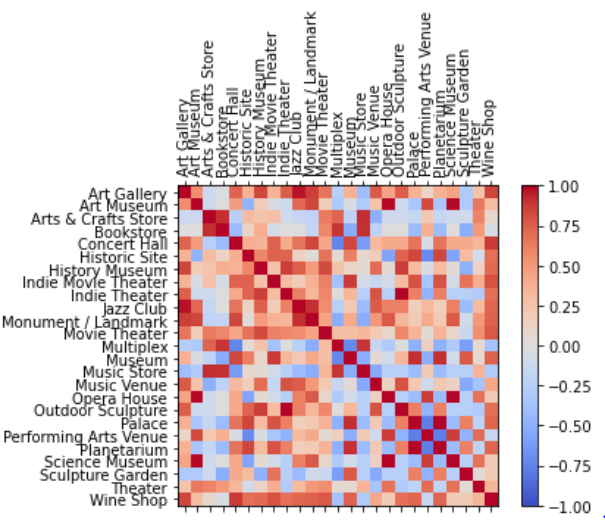
**Data**

As the budget for this study was limited, the data scientist chose to use publicly available data sources. The first of these is the list of London Boroughs sourced from Wikipedia, providing location and population size. The next dataset is derived from Foursquare.com, providing information on existing venues in each Borough. The venue dataset includes information on the number of current entertainment venues and their distribution around London, which is an important metric to predict where the art gallery will be most popular.

**Methodology**

Location data for each Borough was first gathered from Wikipedia. This data was cleaned to remove empty and non-useful columns. The latitude and longitude from this data was then used search Foursquare.com for venues in a 5km radius from the centre of each Borough. Preliminary analysis needed to take place to understand whether there were any relationships that could be drawn between different venue types or Boroughs and as such a correlation matric was used. Additionally, statistical analysis was carried out looking at the mean, sum and count of each venue type per Borough and finally listing the top 10 types of venue in each Borough. This enabled an understanding of which venues tended to be in the same place as Art Galleries.

Tools used to understand the correlation between different venue types included using scatter plots, linear regression and correlation matrices – example below demonstrating the relationships between artistic venues (figure 1):



*Figure 1. Correlation matrix between different artistic venues in London, red indicating a positive relationship and blue a negative relationship*

The correlations between individual parameters were very variable and therefore it was decided to run a clustering machine algorithm, k-means, to group the Boroughs based on multiple parameters. This algorithm partitions the different Boroughs into clusters, based on the commonality of venues, in which each Borough belongs to the cluster with the nearest mean. Essentially this algorithm allows the similarities of Boroughs in the same cluster to be highlighted and therefore inferences to be drawn. The k-means number was varied to optimise the clusters and resulted in separating the Boroughs into 5 different clusters.

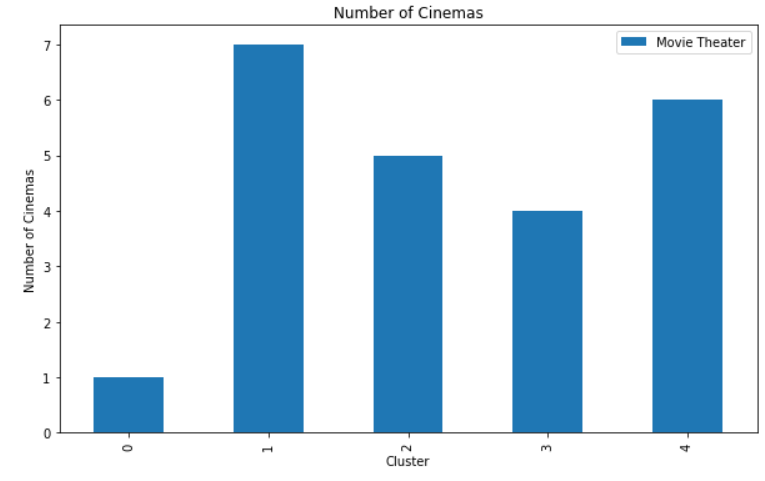
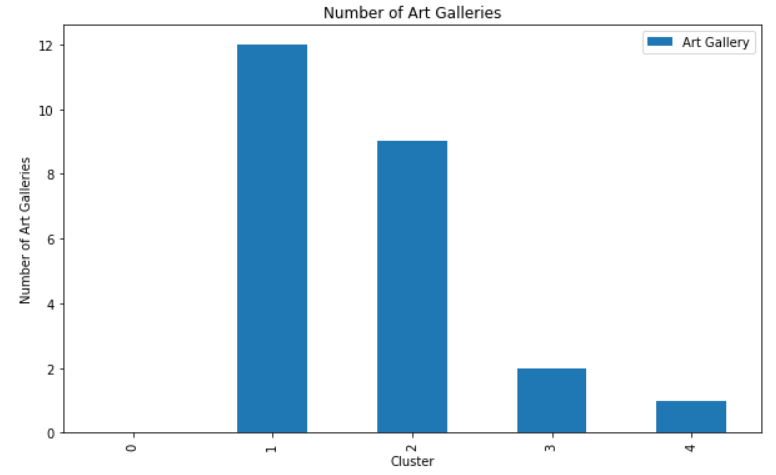
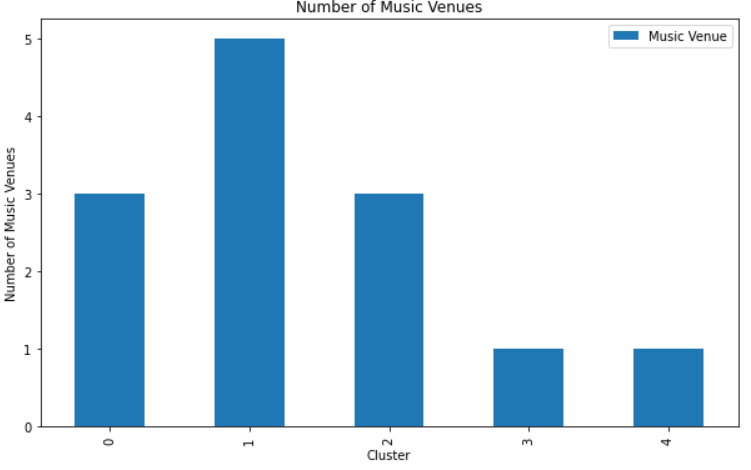
The next step was to identify which Boroughs currently had the most entertainment venues as this could be used as a potential indicator for the popularity in an art gallery. Any Boroughs without an art gallery in this cluster could be an opportunity.

**Results**

Analysis of the results indicated that the most common venues in each cluster were as follows:

* Cluster 0: Indian restaurants, coffee shops and gyms
* Cluster 1: Parks and art galleries
* Cluster 2: Hotels
* Cluster 3: Pubs
* Cluster 4: Pubs, parks and supermarkets

Deeper analysis plotting the number of entertainment venues (art galleries, music venues and movie theaters) in each cluster demonstrated that these venue types where highest in cluster 1 (see figure 2).

,  

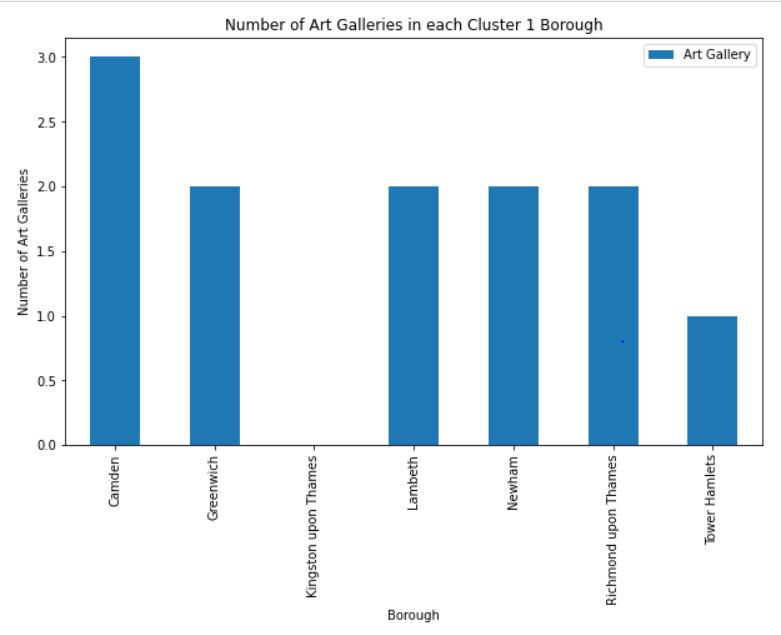
*Figure 2c. Bar chart demonstrating that Cluster 1 also has the greatest number of Music Venues.*

*Figure 2b. Bar chart demonstrating that Cluster 1 also has the greatest number of Movie Theaters.*

*Figure 2a. Bar chart demonstrating that Cluster 1 has the greatest number of Art Galleries.*

**Discussion**

The results show that Boroughs in cluster 1 contain the highest number of art galleries as well as similar venues. This suggests that the local population in these areas may be interested in art galleries and that these would be low risk locations to open the art gallery. All Boroughs in cluster 1 contain an art gallery, except ‘Kingston-upon-Thames’ (figure 3). Therefore, it is recommended to open the new art gallery in Kingston-upon-Thames as there will be no competitors, with Tower Hamlets being the next best option with only one existing art gallery.



*Figure 2c. Bar chart demonstrating that Cluster 1 also has the greatest number of Music Venues.*

**Conclusions**

The new art gallery should be set-up in Kingston-upon-Thames. If no premises are available there, then Tower Hamlets would be the next most suitable.

**References**

Wikipedia: List of London Boroughs <https://en.wikipedia.org/wiki/List_of_London_boroughs#:~:text=List%20of%20boroughs%20and%20local%20authorities%20%20,%20%2016.70%20%207%20more%20rows%20>

Foursquare: Venue information