ESTABLISHING NEW CINEMA LOCATIONS IN LONDON, UK

COURSERA CAPSTONE – IBM DATA SCIENCE PROFESSIONAL CERTIFICATE

BRIAN COLLINS

MAY 2021

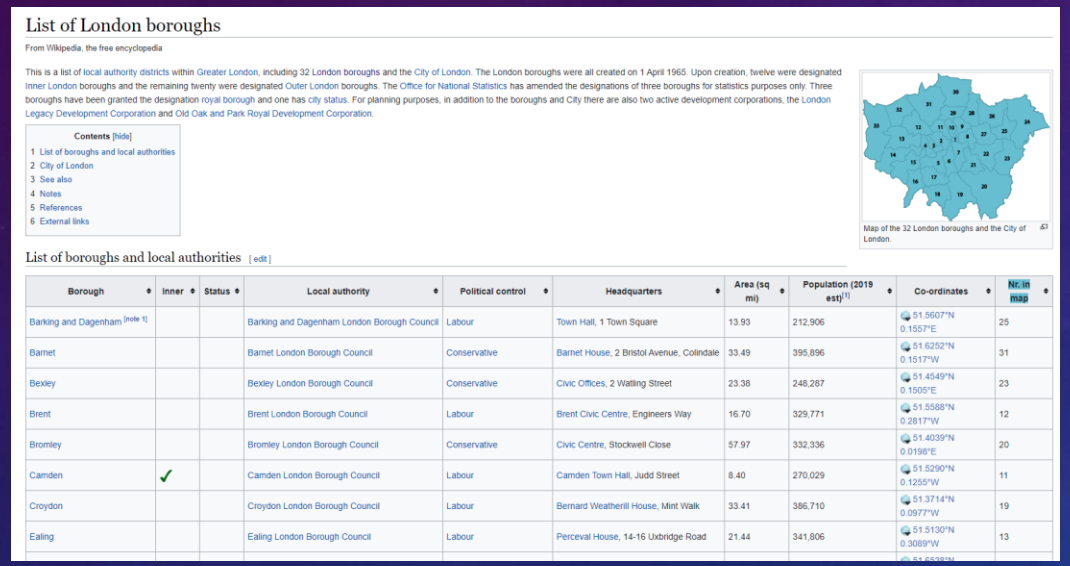
CONCEPT

- Ever since the UK government introduced new planning policy guidelines aimed at revitalising city centres and control out-of-town developments, planners and developers have seen cinemas as key amenities to make urban masterplans and their associated schemes more attractive.
- As major retailers come under pressure to attract shoppers with leisure-related activities, a growing number of cinema construction projects are being built across the country to help ensure that new developments succeed.
- The focus of this project will be establishing which Boroughs in the Greater London Area are underserved by both established cinema chains alongside smaller independent cinemas.

DATA SOURCES

- Office of National Statistics population data will be used to establish the population levels within each of the 32 London boroughs.
- Wikipedia entry for the London Boroughs will be utilised to gain an oversight of the names of both Boroughs and Local Authorities that control those Boroughs
- Utilising the Foursquare API to get an understanding of the number of cinemas already present within each Borough currently.

DATA SCRAPED FROM WIKIPEDIA



List of London boroughs

From Wikipedia, the free encyclopedia

This is a list of local authority districts within Greater London, including 32 London boroughs and the City of London. The London boroughs were all created on 1 April 1965. Upon creation, twelve were designated Inner London boroughs and the remaining twenty were designated Outer London boroughs. The Office for National Statistics has amended the designations of three boroughs for statistics purposes only. Three boroughs have been granted the designation royal borough and one has city status. For planning purposes, in addition to the boroughs and City there are also two active development corporations, the London Legacy Development Corporation and Old Oak and Park Royal Development Corporation.

Contents [hide]

1 List of boroughs and local authorities
2 City of London
3 See also
4 Notes
5 References
6 External links

Map of the 32 London boroughs and the City of London.

List of boroughs and local authorities [edit]

Borough	Inner	Status	Local authority	Political control	Headquarters	Area (sq mi)	Population (2019 est)[1]	Co-ordinates	Nr. in map
Barking and Dagenham [note 1]			Barking and Dagenham London Borough Council	Labour	Town Hall, 1 Town Square	13.93	212,906	51.5607°N 0.1557°E	25
Barnet			Barnet London Borough Council	Conservative	Barnet House, 2 Bristol Avenue, Colindale	33.49	395,896	51.6252°N 0.1517°W	31
Bexley			Bexley London Borough Council	Conservative	Civic Offices, 2 Watling Street	23.38	248,287	51.4549°N 0.1505°E	23
Brent			Brent London Borough Council	Labour	Brent Civic Centre, Engineers Way	16.70	329,771	51.5588°N 0.2817°W	12
Bromley			Bromley London Borough Council	Conservative	Civic Centre, Stockwell Close	57.97	332,336	51.4039°N 0.0198°E	20
Camden	✓		Camden London Borough Council	Labour	Camden Town Hall, Judd Street	8.40	270,029	51.5290°N 0.1255°W	11
Croydon			Croydon London Borough Council	Labour	Bernard Weatherill House, Mint Walk	33.41	386,710	51.3714°N 0.0977°W	19
Ealing			Ealing London Borough Council	Labour	Perceval House, 14-16 Uxbridge Road	21.44	341,806	51.5130°N 0.3089°W	13

- To produce a list of each of the London Boroughs, in addition to further information about their political control, data was scraped using from a table on Wikipedia using Urllib.
- Once the data was scraped and assigned to a variable, we used Beautiful Soup to extract and work with the data in it.
- The class ID of the important tables on the page was 'Wikitable sortable' so a further function pulled out only tables with this ID
- Looped through the rows of the table using the tr (row) and td (cell) fields to populate the a Dataframe.

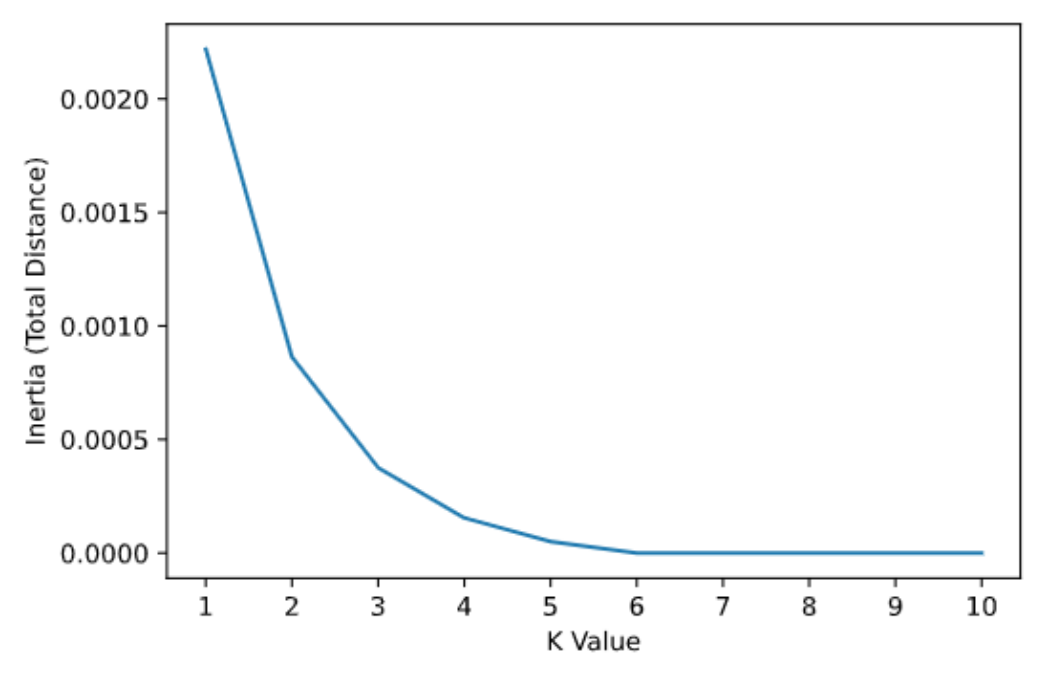
ADDING GEOGRAPHIC COORDINATES

- Once a table had been created and stored the data for each of the London Boroughs Geocoder was used to find the Latitude and Longitude of each of the geographic centres of the Boroughs.
- Source table did include Lat/Long information, this was for the political headquarters of each Borough, rather than the centre, the latter being most suitable for the analysis. This was saved as a separate Dataframe.
- Thus, Geocoder was used to find the Latitude and Longitude of each of the geographic centres of the Boroughs.

USING THE FOURSQUARE API

- These geographic coordinates were then passed through the Foursquare API to look for nearby venues within a 6 Kilometre range from the centre of each borough.
- A onehot was then produced to establish the mean number of each venue category produced by the resultant search and this Dataframe was then queried for unique values within 'Venue Category’
- This established 'Movie Theatre' and 'Indie Movie Theatre' as the categories of interest for this study. Onehots for these categories were extracted for each borough and saved as a new Dataframe.

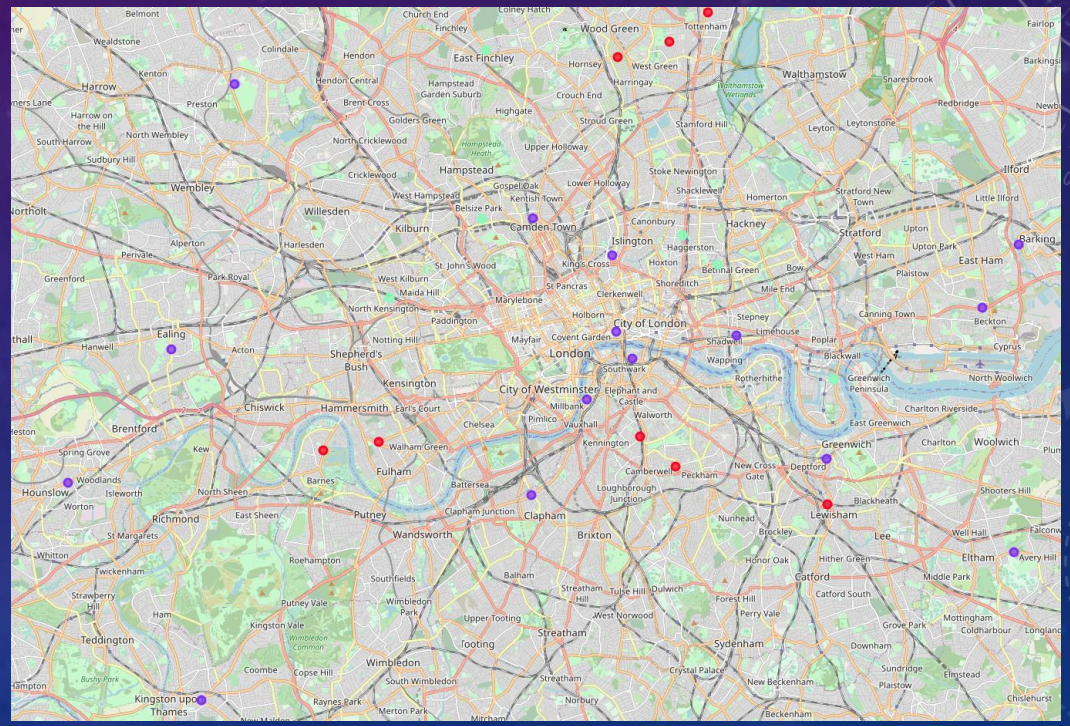
K-MEANS – CLUSTERING USING MACHINE LEARNING



- K-means Clustering was then utilised to cluster boroughs based on the output around the venue categories previously established.
- Prior to a full run, a Cluster Variance function was defined to establish inertia looking at total Euclidean distance between data values for K=1 through K=10.
- A K value of 3 was ultimately used as this produced three clear clusters, one with presence of either a Movie Theatre, an Independent Movie Theatre or neither.

K-MEANS CLUSTERS

- Three clusters divided the dataset into three distinct categories.
    - Boroughs well served by independent cinemas without any real coverage of mainstream chains and the influence brought by size of location, both in terms of number of screens and maximum footfall.
    - Boroughs with good coverage of mainstream cinemas, without a great deal of coverage of independent locations, and the ability for those locations to show the mixture of blockbuster and arthouse films that make them so well loved by their clientele.
    - Boroughs that are completely unserved by either type of location and might possibly present good investment opportunities by developers or chains looking to open new locations. This however leads on to the next analysis.



CONSIDERING POPULATION DATA

```
[59]
no_theater = london_merged.loc[london_merged['Cluster Labels'] == 1]
borough_joined = pd.merge(borough_table, no_theater, on='Borough', how='inner')
borough_joined.drop(['Indie Movie Theater', 'Movie Theater'], axis=1, inplace=True)
borough_joined.sort_values(['Population (2019 est)'], ascending=False, inplace=True)
borough_joined
```

	Borough	Local authority	Political control	Area (sq mi)	Population (2019 est)	Latitude_x	Longitude_x	Cluster Labels
10	Newham	Newham London Borough Council	Labour	13.98	353,134	51.5077	0.0469	1
2	Ealing	Ealing London Borough Council	Labour	21.44	341,806	51.5130	-0.3089	1
14	Wandsworth	Wandsworth London Borough Council	Conservative	13.23	329,677	51.4567	-0.1910	1
9	Lambeth	Lambeth London Borough Council	Labour	10.36	326,034	51.4607	-0.1163	1
11	Southwark	Southwark London Borough Council	Labour	11.14	318,830	51.5035	-0.0804	1
3	Greenwich	Greenwich London Borough Council	Labour	18.28	287,942	51.4892	0.0648	1
13	Waltham Forest	Waltham Forest London Borough Council	Labour	14.99	276,983	51.5908	-0.0134	1
6	Hounslow	Hounslow London Borough Council	Labour	21.61	271,523	51.4746	-0.3680	1
15	Westminster	Westminster City Council	Conservative	8.29	261,317	51.4973	-0.1372	1
5	Havering	Havering London Borough Council	Conservative	43.35	259,552	51.5812	0.1837	1
4	Harrow	Harrow London Borough Council	Labour	19.49	251,160	51.5898	-0.3346	1
1	Bexley	Bexley London Borough Council	Conservative	23.38	248,287	51.4549	0.1505	1
7	Islington	Islington London Borough Council	Labour	5.74	242,467	51.5416	-0.1022	1
0	Barking and Dagenham	Barking and Dagenham London Borough Council	Labour	13.93	212,906	51.5607	0.1557	1
12	Sutton	Sutton London Borough Council	Liberal Democrat	16.93	206,349	51.3618	-0.1945	1
8	Kingston upon Thames	Kingston upon Thames London Borough Council	Liberal Democrat	14.38	177,507	51.4085	-0.3064	1

- Arguably the most significant factor in deciding where to open a new cinema might be the potential number of customers that would help any newly opened location reach the critical mass of business required to transition into an established site.
- Overlaid ONS population data on cluster group completely unserved by either major chains or independent cinemas.
- This indicated that locations such as Newham or Ealing are best placed as areas of further exploration.

LIMITATIONS OF STUDY

Geographic coverage

- Searches cantered around a single geographical point for each of the boroughs with a set radius leaves open the possibility that a particular location, geographically proximate to two of the centres of each borough could well have been identified twice
- Conversely, it is entirely possible that a location was outside of the 6-kilometre radius that represented the search zone of venues around those Borough centroids.
- Worthwhile expanding the search radius, but then completing extra work to remove duplicates from the study. Perhaps assigning a location identified multiple times to the centroid that it is closest to, rather than either counting multiple times, or not at all.

Comparing potential locations

- Reasonable to suggest that any of the top 5 or 10 locations would make suitable candidates for a new cinema location.
- Therefore useful to conduct further analysis on additional complications such as average travel times to and from each location to establish the ease of access to any new site using existing public transport links.

IN SUMMARY

- It stands to reason that markets are poised to return to some semblance of normality in a post-Covid world.
- Businesses are certainly planning their exit strategy which may well involve expanding their footprint to include additional sites in underserved communities.
- It is with this fact in mind that this report focused on the integration of data science and machine learning technics to assist decision makers in making informed and efficient choices.
- This study was narrow in scope and certainly would require a significant expansion to provide any results that would be actionable in a real-world scenario.
- It could be reasonable to state that from the results provided, we could potentially recommend potential locations for future research into their theoretical profitability as venue locations.