Best boroughs in London to live it

Project: IBM Capstone
By: Umair Arif
Date; 24th December 2020

1-Introduction

No London borough Is quite the same as another. There are 32 boroughs in total, they all have different attractions, architecture styles and atmospheres which makes them feel almost like a small cluster of different towns and cities.

We have been approached by a magazine, who wants to publish an article on the best boroughs to live in their upcoming edition. They have tasked us to determine the quality of life in each borough and the most ideal boroughs for outdoor activities, shopping destinations, nightlife and dinning.

In this report we will extract the data on the transport accessibility, happiness score, pollution, crime rates and availability of green spaces in each borough to determine the quality of living, followed by using Foursquare-API to construct a model to determine the most popular venues in each boroughs.

2-Data

The data needed for this project is extracted from three databases i.e. the Greater London Authority Datastore, Wikipedia and Foursquare API. Once the data from these databases is imported on python interface, they will undergo the process of formatting and cleaning to remove unwanted elements and to prepare the DataFrame for in depth analysis

2.1Quality of life analysis

The data required for quality of life analysis is obtained from Greater London Authority Datastore, which consists of data regarding the happiness score, transport accessibility, pollution, crime rates and greenspace available in each borough. The data can be downloaded for the Datastore as an Excel format file, which can then be opened on python interface using pandas.

	BoroughName	Crime rate per thousand	% of greenspace	Total carbon emission	Transport accessibility	Happiness score
0	Barking and Dagenham	83.359060	33.6	643.788244	2.970626	7.05
1	Barnet	62.738426	41.3	1415.107370	2.996701	7.37
2	Bexley	51.827942	31.7	974.621259	2.552134	7.21
3	Brent	78.801881	21.9	1175.073235	3.653713	7.22
4	Bromley	64.133577	57.8	1179.800587	2.779859	7.44
5	Camden	123.498709	24.8	1288.317870	5.690269	7.11
6	Croydon	77.035133	37.1	1237.287433	3.224080	7.18

Figure 1: Shows how the data will look on Jupyter Notebook.

2.2 Population and geographical coordinates

The geographical coordinates and the population for each borough is extracted from Wikipedia page. The data is loaded on python using Beautifulsoup. Once the data is loaded on the Jupyter Notebook, the DataFrame will undergo cleaning process to remove unwanted columns and texts such as notes annotation, to split the geographical coordinates in two columns namely, latitudes and longitudes.

Figure 2: raw data from Wikipedia

	BoroughName	Population	Latitude	Longitude
0	Barking and Dagenham	194352	51.5607	0.1557
1	Barnet	369088	51.6252	-0.1517
2	Bexley	236687	51.4549	0.1505
3	Brent	317264	51.5588	-0.2817
4	Bromley	317899	51.4039	0.0198

Figure 3: Data after undergoing formatting and cleaning.

2.3 Merging the DataFrame

Once both the tables are cleaned and filtered, both the tables will be merged to form one database

	BoroughName	Crime rate per thousand	% of greenspace	Total carbon emission	Transport accessibility	Happiness score	Population	Latitude	Longitude
0	Barking and Dagenham	83.4	33.6	643.8	3.0	7.0	194352	51.5607	0.1557
1	Barnet	62.7	41.3	1415.1	3.0	7.4	369088	51.6252	-0.1517
2	Bexley	51.8	31.7	974.6	2.6	7.2	236687	51.4549	0.1505
3	Brent	78.8	21.9	1175.1	3.7	7.2	317264	51.5588	-0.2817

Figure 4: Shows the merged DataFrame

2.4 Venues

The venues and venue category data is obtained from Foursquare API. The sample of the data on python interface is shown in figure 5.

	BoroughName	Borough Latitude	Borough Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Barking and Dagenham	51.5607	0.1557	Central Park	51.559560	0.161981	Park
1	Barking and Dagenham	51.5607	0.1557	Crowlands Heath Golf Course	51.562457	0.155818	Golf Course
2	Barking and Dagenham	51.5607	0.1557	Robert Clack Leisure Centre	51.560808	0.152704	Martial Arts School

Figure 5: sample of what the data extracted from Foursquare looks like.

Borough	Inner	Status	Local authority	Political control	Headquarters	Area (sq mi)	Population (2013 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	194,352	51.5607°N 0.1557°E	25
Barnet			Barnet London Borough Council	Conservative	Barnet House, 2 Bristol Avenue, Colindale	33.49	369,088	51.6252°N 0.1517°W	31
Bexley			Bexley London Borough Council	Conservative	Civic Offices, 2 Watling Street	23.38	236,687	51.4549°N 0.1505°E	23