Finding a house in London

Introduction

Mr. and Mrs. Castro want to live in London, he is from Mexico and she is from the USA, they are in their 20s, they are recently married, they have no children and they do not plan to have them until after 10 years.

A month ago they were on their honeymoon and visited London, and fell in love with the city, so much so that they decided to move to this great and exciting city.

How not to fall in love with London, London is a great city, with a lot of life and fascinating corners to discover that are distributed in the 33 districts that make up this city. The most outstanding feature of London is that it is a city with a lot of personality, capable of making travelers from all over the world fall in love with just a glance. When you visit it, you will be able to see many emblematic buildings with history and an architectural level that will leave you ecstatic. In addition, you will find in it numerous alternatives for fun, activities of all kinds, nightlife.

As we can see, because of their backgrounds they know very little about the Borough in London. So I decided to help them.

According to their lifestyles, what would be the best group of Borough to live for them?

Considering that they would have to sell their house in Los Angeles and that for this they would get a maximum of 500,000 British Pounds, and with this they would have to buy their new home, in addition, it will be necessary to consider that they like crowded neighborhoods,where the most popular are restaurants of all kinds of food and bars, since they do not want to lose all the fun that this multicultural city offers.

Methodology

We will use the database consulted from Wikipedia as the data source, List of London Boroughs.

1. Wikipedia list of Boroughs with their coordinates.
2. Base data on average prices house in london.
3. Foursquare data to the most populars venues.

The source base with which we will answer this question is extracted directly from Wikipedia and, as expected, it will come with some irregularities for what we need.

```
/usr/local/lib/python3.6/dist-packages/statsmodels/tools/_testing.py:19: FutureWarning: pandas.util.testing is deprecated. Use the functions in the public API at pandas.testing instead
  import pandas.util.testing as tm
```

	Borough	Inner	Status	Local authority	Political control	Headquarters	Area (sq mi)	Population (2013 est)[1]	Co-ordinates	Nr. in map
0	Barking and Dagenham [note 1]	NaN	NaN	Barking and Dagenham London Borough Council	Labour	Town Hall, 1 Town Square	13.93	194352	51°33′39″N 0°09′21″E / 51.5607°N 0.1557°E	25
1	Barnet	NaN	NaN	Barnet London Borough Council	Conservative	Barnet House, 2 Bristol Avenue, Colindale	33.49	369088	51°37′31″N 0°09′06″W / 51.6252°N 0.1517°W	31
2	Bexley	NaN	NaN	Bexley London Borough Council	Conservative	Civic Offices, 2 Watling Street	23.38	236687	51°27′18″N 0°09′02″E / 51.4549°N 0.1505°E	23
3	Brent	NaN	NaN	Brent London Borough Council	Labour	Brent Civic Centre, Engineers Way	16.70	317264	51°33′32″N 0°16′54″W / 51.5588°N 0.2817°W	12
4	Bromley	NaN	NaN	Bromley London Borough Council	Conservative	Civic Centre, Stockwell Close	57.97	317899	51°24′14″N 0°01′11″E / 51.4039°N 0.0198°E	20

So we start by removing the columns that are too much in the source using the dorp code.

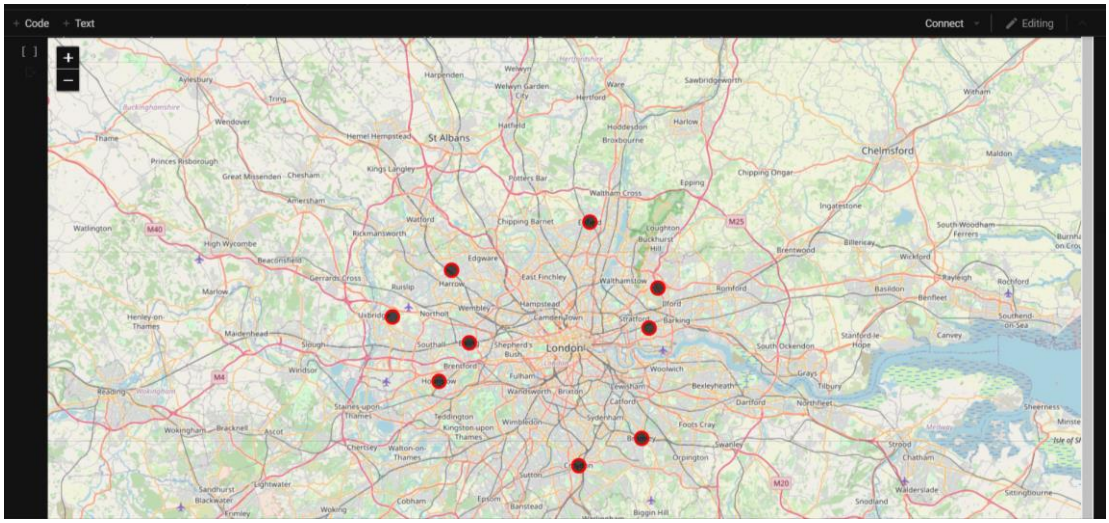
We also skew the information that we are going to require since we do not want the interior Boroughs because our young couple definitely wants to live where things happen.

Finally we proceed to fix the names of the Boroughs that are found with some more brands.

With this process we will get something like this:

	Borough	Area	Population	Co-ordinates
0	Barking and Dagenham	13.93	194352	51°33′39″N 0°09′21″E / 51.5607°N 0.1557°E
1	Barnet	33.49	369088	51°37′31″N 0°09′06″W / 51.6252°N 0.1517°W
2	Bexley	23.38	236687	51°27′18″N 0°09′02″E / 51.4549°N 0.1505°E
3	Brent	16.70	317264	51°33′32″N 0°16′54″W / 51.5588°N 0.2817°W
4	Bromley	57.97	317899	51°24′14″N 0°01′11″E / 51.4039°N 0.0198°E

Finally we are ready to see the Boroughs of London on the map.

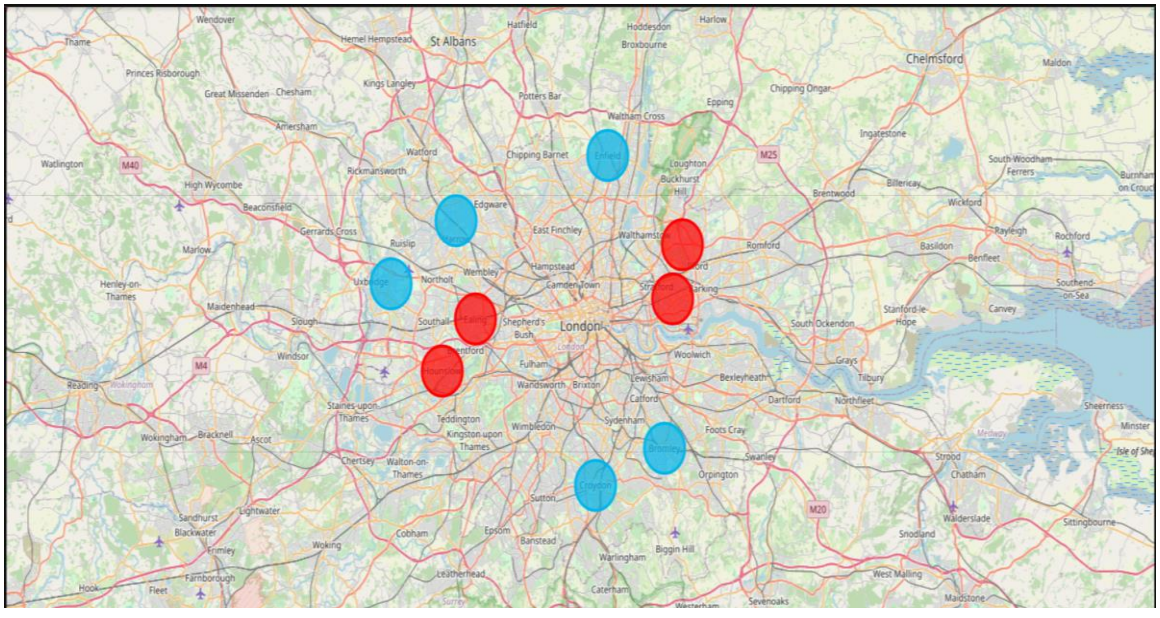


As the idea is to discover in which Borough in London this young couple should go to look for houses in order to move, we will have to consult the Foursquare data source that will give us the list of the most visited places by Boroughs.

	Borough	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Bexley	Pizza Place	Coffee Shop	Ice Cream Shop	Brewery	Sandwich Place
1	Brent	Fast Food Restaurant	Gas Station	Sandwich Place	Convenience Store	Pizza Place
2	Bromley	Coffee Shop	Pub	Grocery Store	Park	Pizza Place
3	Croydon	Pub	Park	Coffee Shop	Grocery Store	Pizza Place
4	Ealing	Park	Pub	Coffee Shop	Pizza Place	Café
5	Enfield	Coffee Shop	Pub	Turkish Restaurant	Supermarket	Café
6	Harrow	Indian Restaurant	Coffee Shop	Pub	Café	Park
7	Havering	Hotel	Thrift / Vintage Store	Food & Drink Shop	Electronics Store	Dessert Shop
8	Hillingdon	Pub	Supermarket	Coffee Shop	Indian Restaurant	Burger Joint
9	Hounslow	Park	Pub	Garden	Coffee Shop	Indian Restaurant

Once we have selected and cleaned the data that we require for the analysis, we use the K - Means Clustering method, an unsupervised algorithm based on the similarity of the interest groups.
What this algorithm does is to divide the Boroughs into groups that have similar characteristics.

The idea is to minimize the distance of the Boroughs within the group and to maximize the distances between the groups.



Results of the clusters

- The first cluster of Boroughs seems to correspond to quite fun people, they enjoy outdoor places as well as trendy bars, they are also well-visited options of cafes and international restaurants. This cluster is made up of 4 Boroughs, where the first Borough is almost at the top of the proposed housing budget, however it has an interesting number of residents and everything required by our young couple.

- The second cluster is made up of a single Borough, and it is quite strange, being its most visited place a Hotel, followed by food or electronic shops, the average price of housing in this Borough is quite affordable.

- The third cluster also offers several quite interesting options, its most popular places boasting cafes, bars and restaurants. One of the neighborhoods is at the exact top of the maximum budget for housing that this couple would have.

- The fourth cluster proposed by this automatic algorithm corresponds to just one Borough again, this Borough is at the top of the housing budget and the most popular of the Borough is a fast food restaurant followed by the gas station, it does not fall within the requirements that they did to us at the beginning of the project.

- The fifth cluster looks like residential Boroughs, where pizzas and ice cream are the most popular, it is made up of 2 Boroughs.

Discussion

Reviewing the data obtained, we observed certain quite interesting clusters according to the requirements of the young couple. These could be clusters 1 and 3 that are the most suitable for your lifestyle, however in the two clusters there are two Boroughs that reach the top of their budget.

They should check in the field with their real estate agent the possibility of seeing houses in these Boroughs, since it is more likely that they will come across houses much more expensive than the average observed here.

```
[ ] B_merged.loc[B_merged['Cluster Label'] == 0, B_merged.columns[[0,1,2,4] + list(range(5, B_merged.shape[1]))]]
```

	Borough	Area	Population	Longitude	Averagehousepriece	Cluster Label	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
4	Ealing	21.44	342494	-0.305195	485000	0	Park	Pub	Coffee Shop	Pizza Place	Café
9	Hounslow	21.61	262407	-0.361347	395000	0	Park	Pub	Garden	Coffee Shop	Indian Restaurant
10	Newham	13.98	318227	0.029318	390500	0	Park	Pub	Café	Bar	Indian Restaurant
11	Redbridge	21.78	288272	0.045410	420000	0	Pub	Park	Coffee Shop	Café	Restaurant

```
[ ] B_merged.loc[B_merged['Cluster Label'] == 2, B_merged.columns[[0,1,2,4] + list(range(5, B_merged.shape[1]))]]
```

	Borough	Area	Population	Longitude	Averagehousepriece	Cluster Label	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
2	Bromley	57.97	317899	0.014814	440000	2	Coffee Shop	Pub	Grocery Store	Park	Pizza Place
3	Croydon	33.41	372752	-0.101957	365000	2	Pub	Park	Coffee Shop	Grocery Store	Pizza Place
5	Enfield	31.74	320524	-0.081018	402500	2	Coffee Shop	Pub	Turkish Restaurant	Supermarket	Café
6	Harrow	19.49	243372	-0.337275	500000	2	Indian Restaurant	Coffee Shop	Pub	Café	Park
8	Hillingdon	44.67	286806	-0.448335	415000	2	Pub	Supermarket	Coffee Shop	Indian Restaurant	Burger Joint

Conclusion

Our enthusiastic and happy couple has decided to meet with their agent to see homes in the Boroughs of clusters 1 and 3 excluding the two Boroughs where their budget reaches the limit.

With this result you can see very good options according to your tastes in 7 great London Boroughs.

Thanks to the application of this methodology, we have reduced the search options from 33 neighborhoods currently existing in London to just 7 Boroughs, which are the ones that really fit your needs.