SOLVING AN OPTIMAL LOCATION PROBLEM – TO VENTURE INTO A RESTAURANT BUSINESS

RAM SUBRAMANIYAN

In this project we will be determining the optimal location of a business in the London city and Greater London area using;

- Geographical location data
- Foursquare data

Post several assumptions and that only the 'Location', 'Borough', 'Postcode', 'Post-town' will be used for our data analysis. Data is transformed in following views based on location and demographic data.

[21]:

	Location	Borough	Postcode	Post-town
0	Abbey Wood	Bexley, Greenwich	SE2	LONDON
1	Acton	Ealing, Hammersmith and Fulham	W3	LONDON
2	Acton	Ealing, Hammersmith and Fulham	W4	LONDON
3	Angel	Islington	EC1	LONDON
4	Angel	Islington	N1	LONDON

	Local authority	White	Mixed	Asian	Black	Other
22	Lewisham	53.5	7.4	9.3	27.2	2.6
27	Southwark	54.3	6.2	9.4	26.9	3.3
21	Lambeth	57.1	7.6	6.9	25.9	2.4
11	Hackney	54.7	6.4	10.5	23.1	5.3
7	Croydon	55.1	6.6	16.4	20.2	1.8

Using **feature engineering and exploratory data analysis, Perform data analysis for single and multiple neighborhoods;**

Using Four square API data, prepare URL to do API call and obtain data to explore venues of the above said boroughs.

All required Neighborhood, Venues Latitude and Longitude along with Venue category was extracted

[50]:

	Location	Borough	Postcode	Latitude	Longitude
0	Crofton Park	Lewisham	SE4	51.46268	-0.03558
1	Denmark Hill	Southwark	SE5	51.47480	-0.09313
2	Deptford	Lewisham	SE8	51.48114	-0.02467
3	Dulwich	Southwark	SE21	51.44100	-0.08897
4	East Dulwich	Southwark	SE22	51.45256	-0.07076

[87]:

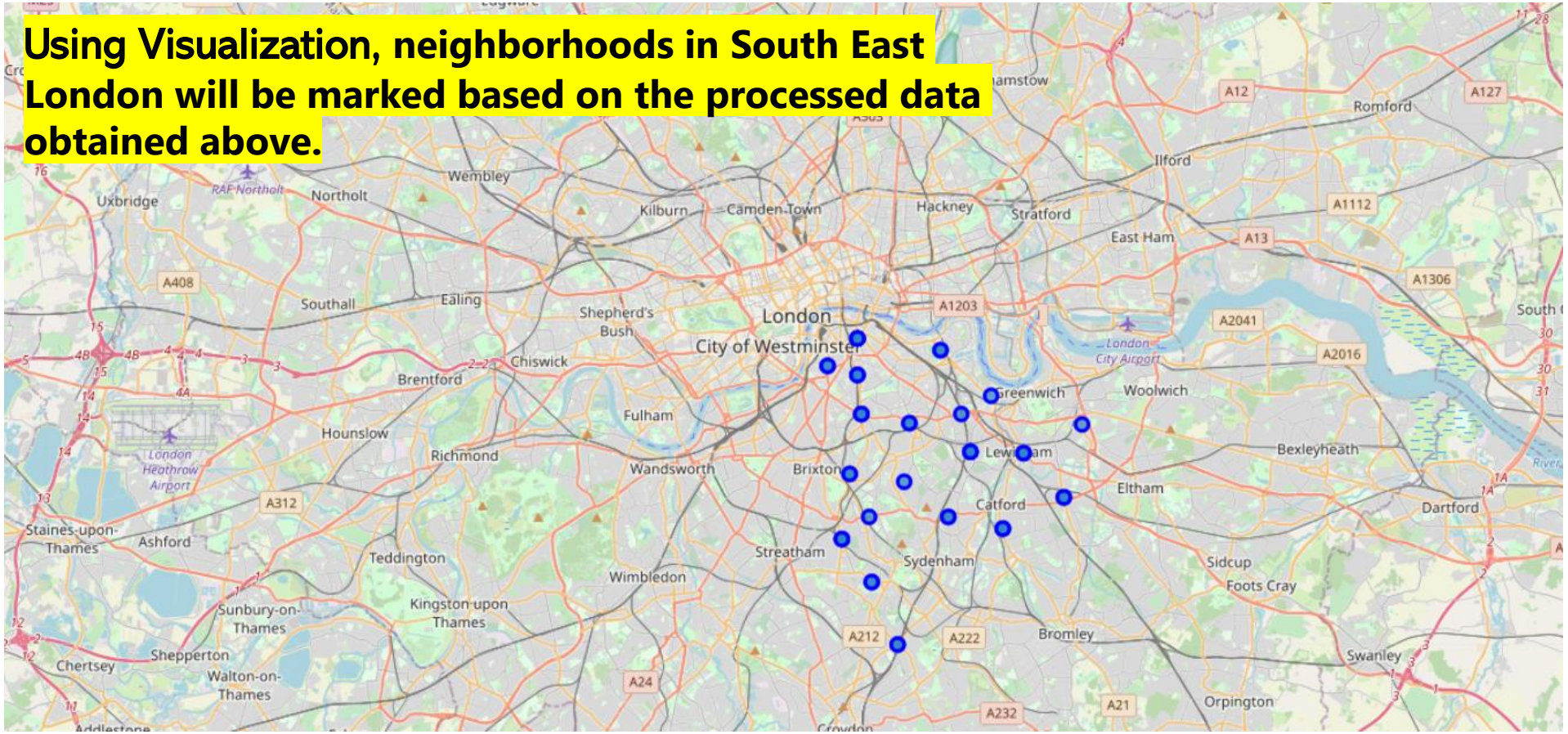
	Count
Pub	413
Coffee Shop	302
Café	273
Park	232
Grocery Store	173

One-hot Encoding

This technique was used to explore the Venue Categories for given neighborhoods. This way, most common venues were determined.

	Neighbourhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Bankside	Hotel	Coffee Shop	Theater	Scenic Lookout	Italian Restaurant	Grocery Store	Cocktail Bar	Pub	Street Food Gathering	Art Museum
1	Bellingham	Grocery Store	Park	Supermarket	Café	Pub	Coffee Shop	Fast Food Restaurant	Italian Restaurant	Train Station	Trail
2	Bermondsey	Hotel	Coffee Shop	Theater	Scenic Lookout	Italian Restaurant	Grocery Store	Cocktail Bar	Pub	Street Food Gathering	Art Museum
3	Blackheath	Pub	Grocery Store	Park	Coffee Shop	Café	Italian Restaurant	Garden	Supermarket	Bakery	Bus Stop
4	Brixton	Café	Coffee Shop	Pub	Park	Italian Restaurant	Middle Eastern Restaurant	Grocery Store	Pizza Place	Cocktail Bar	Bar

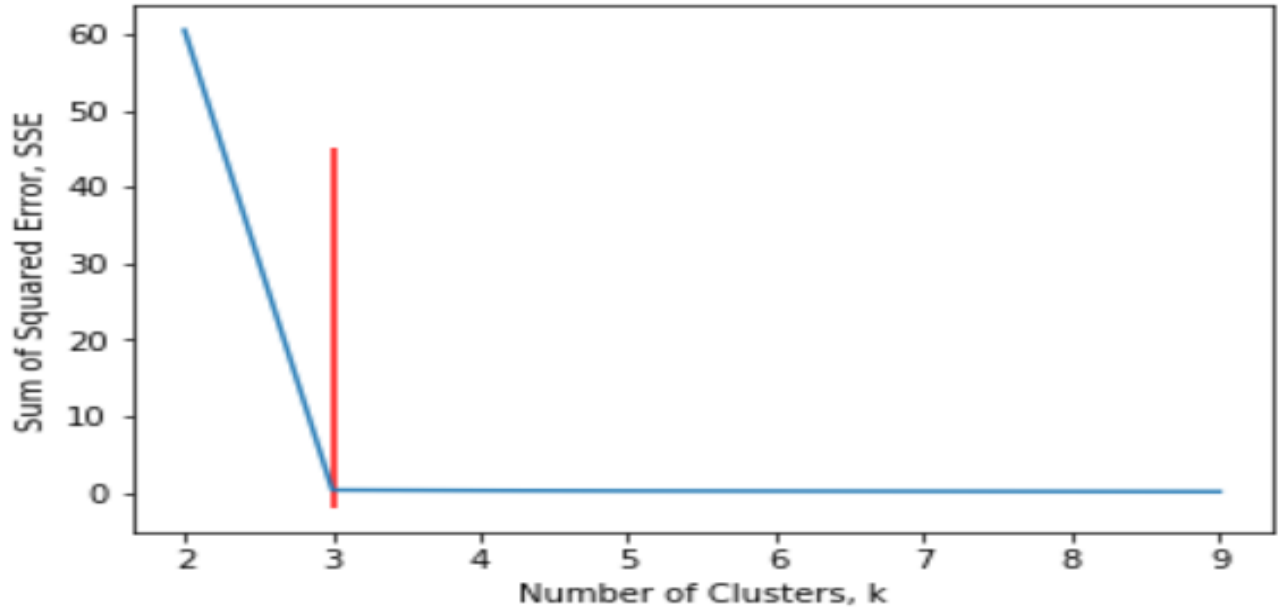
Using Visualization, neighborhoods in South East London will be marked based on the processed data obtained above.



To find the optimal value of the number of clusters, k, the number of clusters is iterated corresponding Silhouette Coefficient is calculated for each of the k-values used. The highest Silhouette Coefficient gives the best match to its own cluster. Please see below:

```
Where n_clusters = 2, the Silhouette Coefficient is 0.5689442041628132
Where n_clusters = 3, the Silhouette Coefficient is 0.6556381843863873
Where n_clusters = 4, the Silhouette Coefficient is 0.6724290452510607
Where n_clusters = 5, the Silhouette Coefficient is 0.7020379193098253
Where n_clusters = 6, the Silhouette Coefficient is 0.7451123840088726
Where n_clusters = 7, the Silhouette Coefficient is 0.835787607388441
Where n_clusters = 8, the Silhouette Coefficient is 0.8600504859117664
Where n_clusters = 9, the Silhouette Coefficient is 0.9345312709880798
```

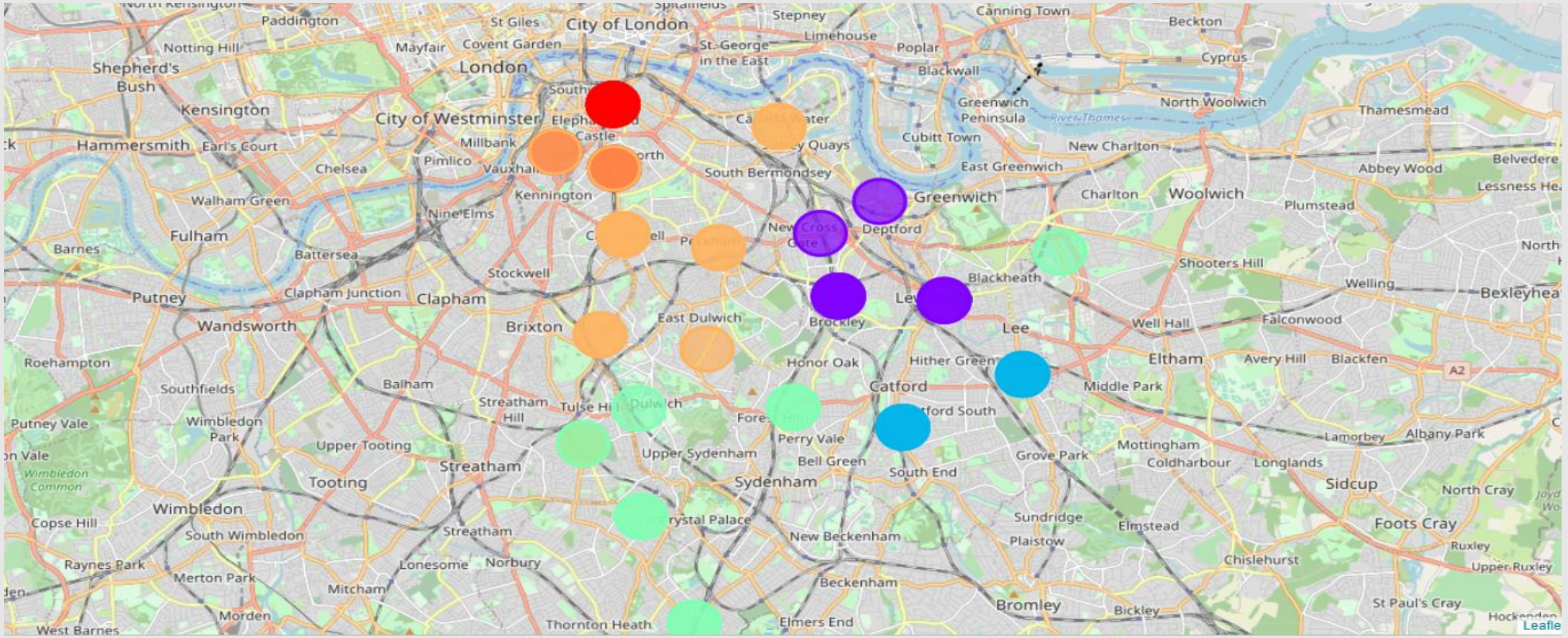
From the result, the high the `n_clusters` the better the silhouette coefficient. For this project, a cluster value of 5 will be used.



Clustering of Neighborhoods

- K-Means clustering technique is used in the project to accomplish goal.
- Generate and checking Cluster labels is necessary to identify what area belongs to which cluster label.
- Using Elbow method and Silhouette Coefficient, no. of clusters to be used for k-means clustering model is determined.
- In this case, 5 Clusters was the optimum.

5 clusters have been shown here on the London map using k-means and evaluation techniques;



	Borough	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Lewisham	1	Pub	Coffee Shop	Café	Park	Indian Restaurant	Italian Restaurant	Cocktail Bar	Bar	Gastropub	Food Truck
2	Lewisham	1	Pub	Coffee Shop	Café	Bar	Park	Gastropub	Garden	Vietnamese Restaurant	Italian Restaurant	History Museum
14	Lewisham	1	Pub	Café	Park	Gastropub	Garden	Coffee Shop	Gym / Fitness Center	Fish & Chips Shop	Restaurant	Food Truck
16	Lewisham	1	Pub	Coffee Shop	Café	Park	Gastropub	Supermarket	Italian Restaurant	Food Truck	Bar	Fish & Chips Shop
17	Lewisham	1	Pub	Coffee Shop	Café	Park	Gastropub	Supermarket	Italian Restaurant	Food Truck	Bar	Fish & Chips Shop
20	Lewisham	1	Pub	Café	Park	Gastropub	Garden	Coffee Shop	Gym / Fitness Center	Fish & Chips Shop	Restaurant	Food Truck
21	Lewisham	1	Pub	Coffee Shop	Café	Bar	Italian Restaurant	Gastropub	Park	Brewery	Indian Restaurant	Vegetarian / Vegan Restaurant
33	Lewisham	1	Pub	Coffee Shop	Café	Park	Indian Restaurant	Italian Restaurant	Cocktail Bar	Bar	Gastropub	Food Truck
42	Lewisham	1	Pub	Coffee Shop	Café	Park	Indian Restaurant	Italian Restaurant	Cocktail Bar	Bar	Gastropub	Food Truck

Of the 5 cluster's data analyzed, using cluster labels assigned in previous steps, it is determined that cluster 2 (cluster label 1) and 3 (cluster label 2) are more suitable in terms of many aspects that reinforces opening a new restaurant / food joint.

	Borough	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
12	Lewisham	2	Grocery Store	Park	Pub	Café	Italian Restaurant	Train Station	Supermarket	Coffee Shop	Gym / Fitness Center	Gym
19	Lewisham	2	Grocery Store	Park	Pub	Café	Italian Restaurant	Train Station	Supermarket	Coffee Shop	Gym / Fitness Center	Gym
26	Lewisham	2	Grocery Store	Park	Supermarket	Café	Pub	Coffee Shop	Fast Food Restaurant	Italian Restaurant	Train Station	Trail
32	Lewisham	2	Grocery Store	Park	Supermarket	Café	Pub	Coffee Shop	Fast Food Restaurant	Italian Restaurant	Train Station	Trail
44	Lewisham	2	Grocery Store	Park	Supermarket	Café	Pub	Coffee Shop	Fast Food Restaurant	Italian Restaurant	Train Station	Trail
45	Lewisham	2	Grocery Store	Park	Pub	Café	Italian Restaurant	Train Station	Supermarket	Coffee Shop	Gym / Fitness Center	Gym

Conclusion

We can conclude with the results above that Cluster 2 and Cluster 3 are the most viable clusters to create a brand multi-cuisine Restaurant. Their proximity to other amenities and accessibility to Train station should be noticed.

These 2 clusters do not have top restaurants that could rival their standards if they are created.

Hence, the most suitable location in Greater London city would be ‘Lewisham’ neighborhood in South East London!!

Thank You!