

Coursera Capstone

The Battle of Neighborhoods (Week 2)

Report Content

1. Introduction section: the business problem to be solved by this project and who may be interested
2. Data selection: Describe data requirements and Sources needed to to solve the problem
3. Methodology section: main component of the report- execute data processing, describe/discuss any data analysis etc.
4. Results section
5. Discussion and Conclusion

INTRODUCTION

The population of London has grown considerably over the last decades. London is very diverse. It represents what is called the reflection of the old British Empire. In London, you can get fresh food supplies from Africa. One begins to wonder how efficient the supply mechanism is.

The real deal is that as much as there are many fine restaurants in London – Asian, Middle Eastern, Latin and American restaurants, you can struggle to find good place to dine in the finest of West African cuisine that has combination of Nigerian, Ghanaian, Cameroonian, Senegalese and more.

My client, a successful restaurant chain in Africa is looking to expand operation into Europe through London. They want to create a high-end restaurant that comes with organic mix and healthy. Their target is not only West Africans, but they are pro-organic and healthy eating. To them every meal counts and counts as a royal when you eat.

Since the London demography is so big, my client needs deeper insight from available data in order to decide where to establish the first Europe “palace” restaurant. This company spends a lot on research and provides customers with data insight into the ingredients used at restaurants. Considering the diversity of London, there is a high multicultural sense. London is a place where different shades live. As such, in the search for an high-end African-inclined restaurant, there is a high shortage. The target audience is broad, it ranges from Londoners, tourists and those who are passionate about organic food

DATA SECTION

This project will rely on public data from Wikipedia and Foursquare.

In this project, London will be used as synonymous to the "Greater London Area". Within the Greater London Area, there are areas that are within the London Area Postcode. The focus of this project will be the neighborhoods are that are within the London Post Code area.

The London Area consists of 32 Boroughs and the "City of London". Our data will be from the link - Greater London Area <https://en.wikipedia.org/wiki/List_of_areas_of_London >

In obtaining the location data of the locations, the Geocoder package is used with the arcgis_geocoder to obtain the latitude and longitude of the needed locations.

The Foursquare API will be used to obtain the South East London Area venues for the geographical location data . These will be used to explore the neighborhoods of London accordingly.

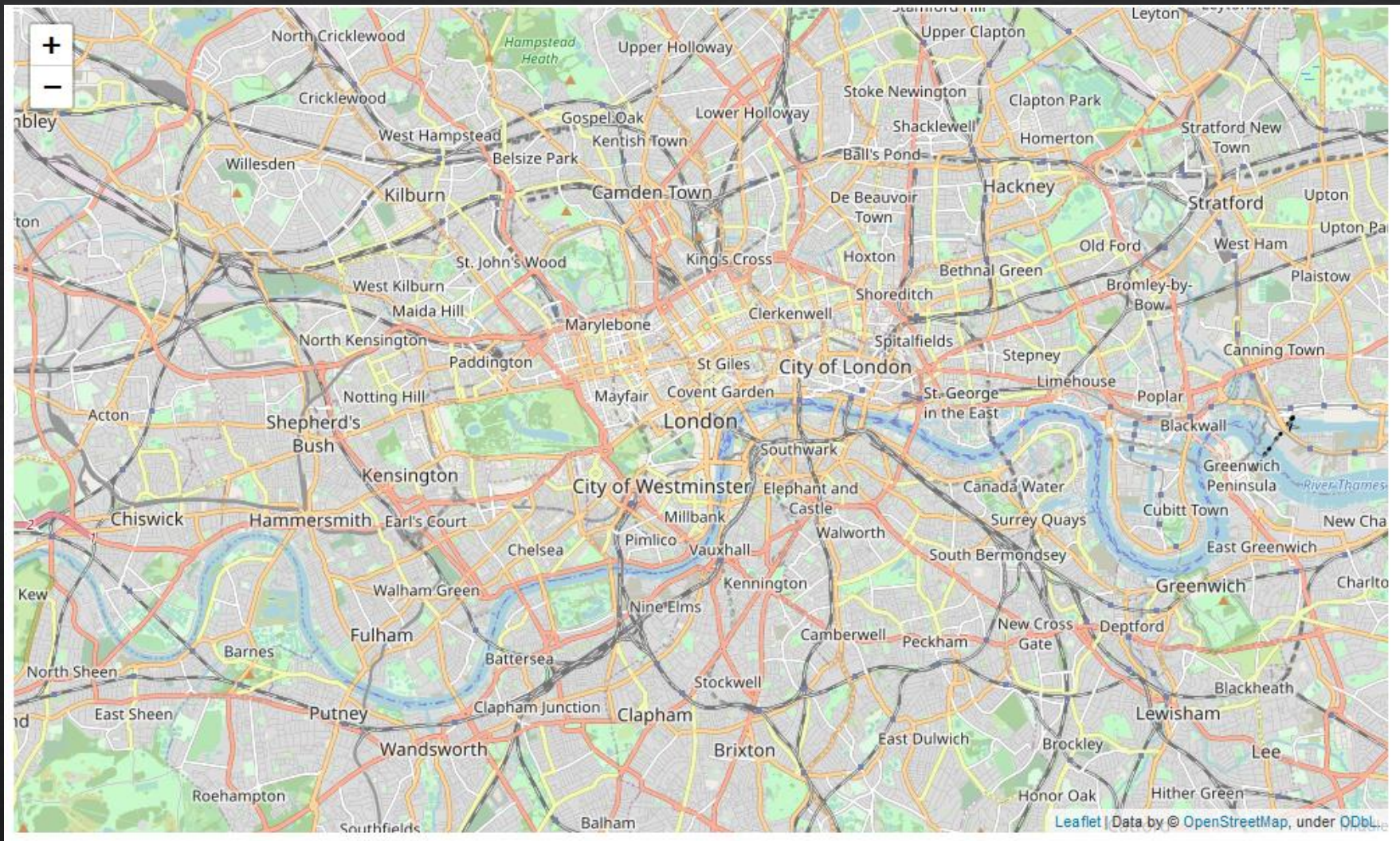
The venues within the neighborhoods of South East London like the areas' restaurants and proximity to amenities would be correlated. Also, accessibility and ease of supplies would be considered as it relates to venues.

METHODOLOGY SECTION

The strategy is based on mapping the describe data in order to facilitate the search of a place to create a brand African Restaurant. The information will be consolidated in one map where one can see the position in detail, the cluster of venues in neighborhood and the relative location from a a subway station for example

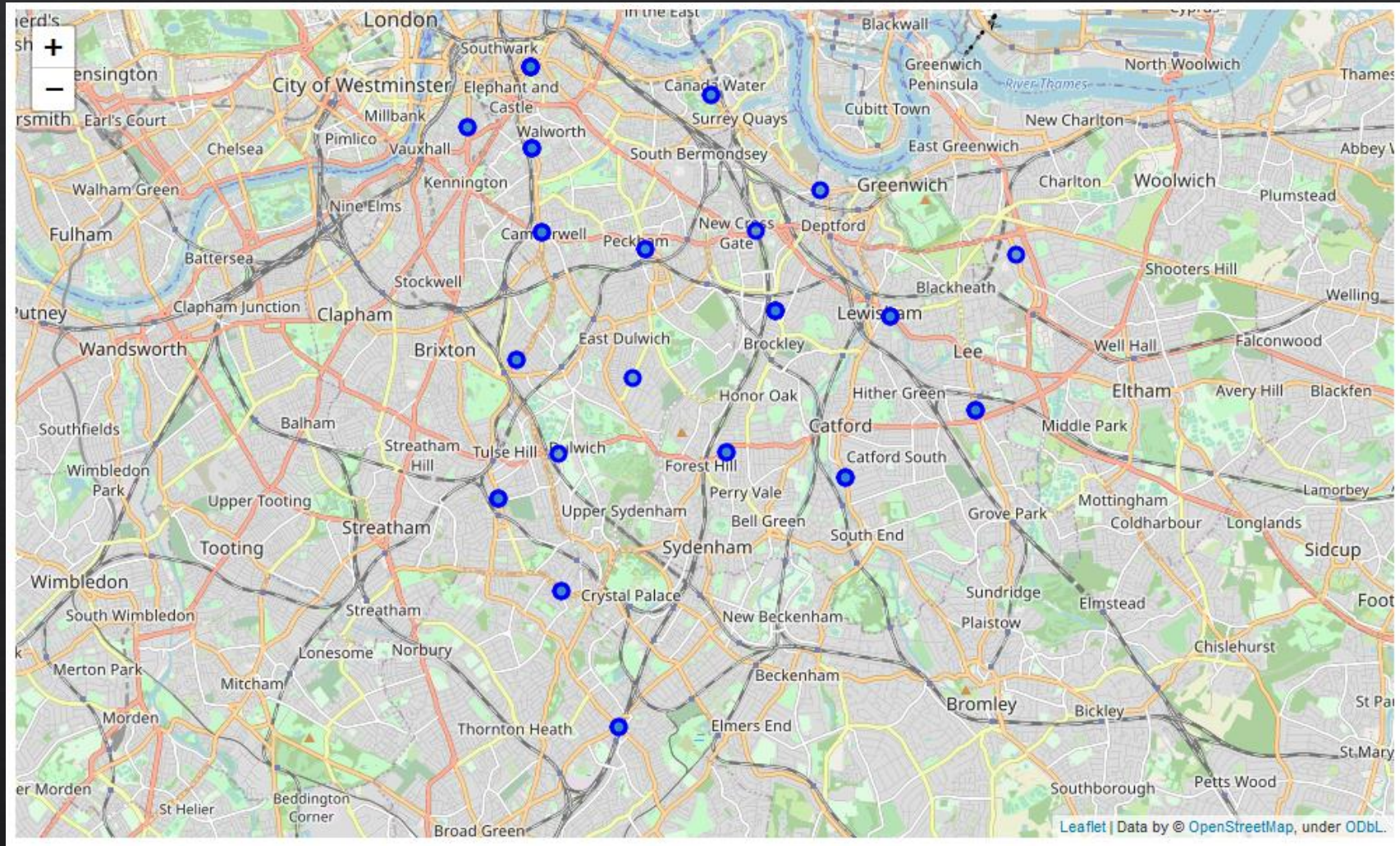
EXECUTION AND RESULTS

LONDON MAP



LONDON MAP

neighborhoods in South East London will be clustered
based on the processed data



Venues in Neighborhood in descending order

	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	Theater	Coffee Shop	Scenic Lookout	Cocktail Bar	Italian Restaurant	Grocery Store	Pub	Street Food Gathering	Art Museum
1	Bellingham	Grocery Store	Supermarket	Park	Coffee Shop	Café	Fast Food Restaurant	Italian Restaurant	Pub	Bus Stop	Clothing Store
2	Bermondsey	Hotel	Theater	Coffee Shop	Scenic Lookout	Cocktail Bar	Italian Restaurant	Grocery Store	Pub	Street Food Gathering	Art Museum
3	Blackheath	Pub	Park	Coffee Shop	Grocery Store	Garden	Italian Restaurant	Supermarket	Bakery	Pizza Place	Clothing Store
4	Brixton	Café	Coffee Shop	Park	Pub	Italian Restaurant	Middle Eastern Restaurant	Bar	Beer Bar	Pizza Place	Cocktail Bar

The next thing to do now, is to create clusters of the neighborhood using the k-means to cluster the neighborhood into 5 clusters

```
In [81]: # set number of clusters
kclusters = 5

# run k-means clustering
kmeans = KMeans(n_clusters = kclusters, random_state=0).fit(se_grouped_clustering)

# check cluster labels generated for each row in the dataframe
kmeans.labels_[0:10]
```

```
Out[81]: array([2, 0, 2, 4, 1, 3, 1, 0, 0, 3], dtype=int32)
```

```
In [82]: kmeans.labels_[0:10]
```

```
Out[82]: array([2, 0, 2, 4, 1, 3, 1, 0, 0, 3], dtype=int32)
```

```
In [83]: # add clustering labels
neighbourhoods_venues_sorted.insert(0, 'Cluster Labels', kmeans.labels_)
```

```
In [84]: se_merged = se_df
```

```
In [85]: se_merged.head(3)
```

```
Out[85]:
```

	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

Merge SE London data with latitude/longitude for each neighborhood

In [87]: `se_merged_latlong.head(5)`

Out [87]:

	Location	Borough	Postcode	Latitude	Longitude	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	
0	Crofton Park	Lewisham	SE4	51.46268	-0.03558	3	Pub	Coffee Shop	Café	Park	Indian Restaurant	Bar	Gastropub	Ita Re
1	Denmark Hill	Southwark	SE5	51.47480	-0.09313	1	Café	Coffee Shop	Park	Pub	Italian Restaurant	Middle Eastern Restaurant	Bar	Be
2	Deptford	Lewisham	SE8	51.48114	-0.02467	3	Pub	Café	Coffee Shop	Park	Bar	Gastropub	Garden	Vi Re
3	Dulwich	Southwark	SE21	51.44100	-0.08897	4	Pub	Café	Park	Grocery Store	Bakery	Coffee Shop	Brewery	Gy Fi Ce
4	East Dulwich	Southwark	SE22	51.45256	-0.07076	1	Café	Park	Pub	Pizza Place	Italian Restaurant	Platform	Burger Joint	Gy Fi Ce

<

>

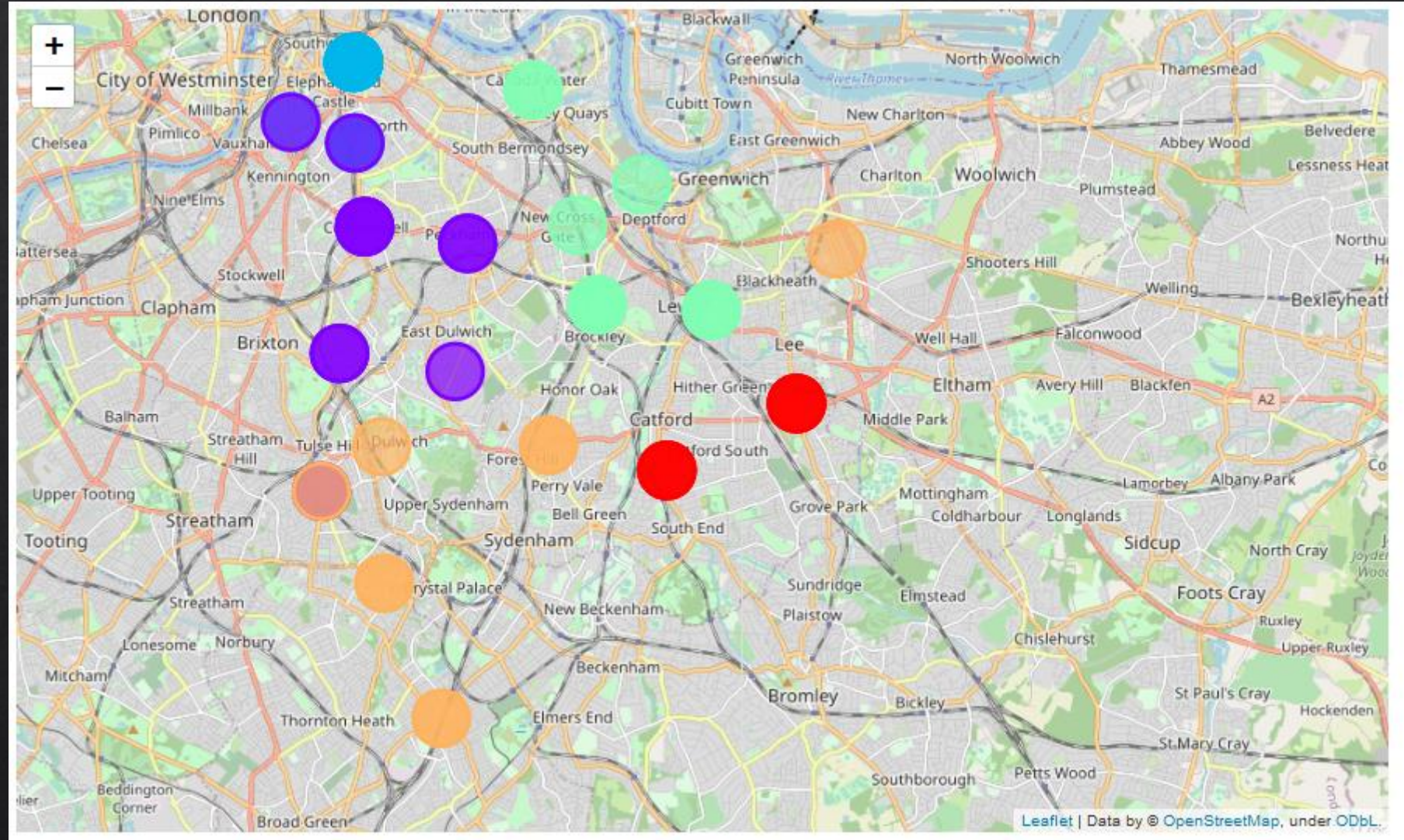
Optimal Number of Clusters for K-mean

To get the optimal number of clusters to be used for the K-mean, there are a number ways possible for the evaluation. Therefore, in this task, the following are used:

The elbow method is used to solve the problem of selecting k . Interestingly, the elbow method is not perfect either, but it gives significant insight that is perhaps not top optimal but sub-optimal to choosing the optimal number of clusters by fitting the model with a range of values for k .

The approach for this is to run the k-means clustering for a range of value k and for each value of k , the Sum of the Squared Errors (SSE) is calculated., calculate sum of squared errors (SSE). When this is done, a plot of k and the corresponding SSEs are then made. At the elbow (just like arm), that is where the optimal value of k is. And that will be the number of clusters to be used. The whole idea is to have minimum SSE. 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

London Map- Neighbourhoods and Cluster of Venues



Cluster 1

	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	0	Grocery Store	Pub	Park	Café	Italian Restaurant	Coffee Shop	Gym / Fitness Center	Supermarket	Train Station	Indian Restaurant
19	Lewisham	0	Grocery Store	Pub	Park	Café	Italian Restaurant	Coffee Shop	Gym / Fitness Center	Supermarket	Train Station	Indian Restaurant
26	Lewisham	0	Grocery Store	Supermarket	Park	Coffee Shop	Café	Fast Food Restaurant	Italian Restaurant	Pub	Bus Stop	Clothing Store
32	Lewisham	0	Grocery Store	Supermarket	Park	Coffee Shop	Café	Fast Food Restaurant	Italian Restaurant	Pub	Bus Stop	Clothing Store
44	Lewisham	0	Grocery Store	Supermarket	Park	Coffee Shop	Café	Fast Food Restaurant	Italian Restaurant	Pub	Bus Stop	Clothing Store
45	Lewisham	0	Grocery Store	Pub	Park	Café	Italian Restaurant	Coffee Shop	Gym / Fitness Center	Supermarket	Train Station	Indian Restaurant

Cluster 2

	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
1	Southwark	1	Café	Coffee Shop	Park	Pub	Italian Restaurant	Middle Eastern Restaurant	Bar	Beer Bar	Pizza Place	Cocktail Bar
4	Southwark	1	Café	Park	Pub	Pizza Place	Italian Restaurant	Platform	Burger Joint	Gym / Fitness Center	Mediterranean Restaurant	Cocktail Bar
13	Lambeth	1	Coffee Shop	Pub	Cocktail Bar	Café	Market	Park	Pizza Place	Burger Joint	Vegetarian / Vegan Restaurant	Indie Movie Theater
24	Southwark	1	Pub	Café	Coffee Shop	Pizza Place	Park	Bar	Gastropub	Cocktail Bar	Art Gallery	Burger Joint
25	Lambeth	1	Café	Pub	Park	Coffee Shop	Hotel	Theater	Italian Restaurant	Gay Bar	Garden	Indian Restaurant
27	Southwark	1	Pub	Café	Coffee Shop	Pizza Place	Park	Bar	Gastropub	Cocktail Bar	Art Gallery	Burger Joint
35	Lambeth	1	Pub	Coffee Shop	Café	Park	Grocery Store	Pizza Place	Bakery	Cocktail Bar	Brewery	Market
36	Lambeth	1	Pub	Coffee Shop	Café	Park	Grocery Store	Pizza Place	Bakery	Cocktail Bar	Brewery	Market
38	Southwark	1	Café	Coffee Shop	Park	Pub	Italian Restaurant	Art Gallery	Pizza Place	Beer Bar	Theater	Cricket Ground
41	Lambeth	1	Café	Coffee Shop	Park	Pub	Italian Restaurant	Middle Eastern Restaurant	Bar	Beer Bar	Pizza Place	Cocktail Bar
43	Southwark	1	Café	Coffee Shop	Park	Pub	Italian Restaurant	Middle Eastern Restaurant	Bar	Beer Bar	Pizza Place	Cocktail Bar

Cluster 3

	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
5	Southwark	2	Café	Coffee Shop	Pub	Hotel	Theater	Italian Restaurant	Park	Grocery Store	Garden	Street Food Gathering
6	Southwark	2	Café	Coffee Shop	Pub	Hotel	Theater	Italian Restaurant	Park	Grocery Store	Garden	Street Food Gathering
7	Southwark	2	Café	Coffee Shop	Pub	Hotel	Theater	Italian Restaurant	Park	Grocery Store	Garden	Street Food Gathering
8	Southwark	2	Hotel	Theater	Coffee Shop	Scenic Lookout	Cocktail Bar	Italian Restaurant	Grocery Store	Pub	Street Food Gathering	Art Museum
18	Lambeth	2	Hotel	Theater	Coffee Shop	Scenic Lookout	Cocktail Bar	Italian Restaurant	Grocery Store	Pub	Street Food Gathering	Art Museum
22	Southwark	2	Coffee Shop	Café	Italian Restaurant	Pub	Hotel	Theater	Park	Cocktail Bar	Grocery Store	Scenic Lookout
23	Southwark	2	Coffee Shop	Café	Italian Restaurant	Pub	Hotel	Theater	Park	Cocktail Bar	Grocery Store	Scenic Lookout
30	Southwark	2	Hotel	Theater	Coffee Shop	Scenic Lookout	Cocktail Bar	Italian Restaurant	Grocery Store	Pub	Street Food Gathering	Art Museum

Cluster 4

	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	3	Pub	Coffee Shop	Café	Park	Indian Restaurant	Bar	Gastropub	Italian Restaurant	Cocktail Bar	Food Truck
2	Lewisham	3	Pub	Café	Coffee Shop	Park	Bar	Gastropub	Garden	Vietnamese Restaurant	Italian Restaurant	Brewery
14	Lewisham	3	Pub	Café	Park	Gastropub	Coffee Shop	Garden	Restaurant	Gym / Fitness Center	Food Truck	Fish & Chips Shop
16	Lewisham	3	Pub	Coffee Shop	Café	Park	Gastropub	Supermarket	Bar	Food Truck	Italian Restaurant	Fish & Chips Shop
17	Lewisham	3	Pub	Coffee Shop	Café	Park	Gastropub	Supermarket	Bar	Food Truck	Italian Restaurant	Fish & Chips Shop
20	Lewisham	3	Pub	Café	Park	Gastropub	Coffee Shop	Garden	Restaurant	Gym / Fitness Center	Food Truck	Fish & Chips Shop
21	Lewisham	3	Pub	Coffee Shop	Café	Italian Restaurant	Bar	Park	Gastropub	Brewery	Indian Restaurant	Beer Store
28	Southwark	3	Pub	Brewery	Park	Coffee Shop	Café	Beer Bar	Bar	Gym / Fitness Center	Food Truck	Vietnamese Restaurant
33	Lewisham	3	Pub	Coffee Shop	Café	Park	Indian Restaurant	Bar	Gastropub	Italian Restaurant	Cocktail Bar	Food Truck
34	Southwark	3	Pub	Brewery	Park	Coffee Shop	Café	Beer Bar	Bar	Gym / Fitness Center	Food Truck	Vietnamese Restaurant
42	Lewisham	3	Pub	Coffee Shop	Café	Park	Indian Restaurant	Bar	Gastropub	Italian Restaurant	Cocktail Bar	Food Truck

Cluster 5

	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
3	Southwark	4	Pub	Café	Park	Grocery Store	Bakery	Coffee Shop	Brewery	Gym / Fitness Center	Farmers Market	Italian Restaurant
9	Lewisham	4	Pub	Grocery Store	Coffee Shop	Café	Park	Gym / Fitness Center	Supermarket	Gastropub	Italian Restaurant	Japanese Restaurant
10	Lambeth	4	Pub	Coffee Shop	Café	Grocery Store	Park	Bakery	Italian Restaurant	Pharmacy	Train Station	Pizza Place
11	Lambeth	4	Pub	Coffee Shop	Café	Grocery Store	Park	Bakery	Italian Restaurant	Pharmacy	Train Station	Pizza Place
15	Lewisham	4	Pub	Grocery Store	Coffee Shop	Café	Park	Gym / Fitness Center	Supermarket	Gastropub	Italian Restaurant	Japanese Restaurant
29	Croydon	4	Pub	Grocery Store	Supermarket	Café	Coffee Shop	Tram Station	Park	Platform	Hotel	Clothing Store
31	Croydon	4	Pub	Grocery Store	Supermarket	Café	Coffee Shop	Tram Station	Park	Platform	Hotel	Clothing Store
37	Croydon	4	Pub	Café	Coffee Shop	Italian Restaurant	Park	Grocery Store	Gastropub	Train Station	Bakery	Breakfast Spot
39	Lewisham	4	Pub	Park	Coffee Shop	Grocery Store	Garden	Italian Restaurant	Supermarket	Bakery	Pizza Place	Clothing Store
40	Lambeth	4	Pub	Grocery Store	Coffee Shop	Café	Park	Bakery	Gym / Fitness Center	Pizza Place	Pharmacy	Portuguese Restaurant

RESULTS

The following are the highlights of the 5 clusters above:

1. Pubs, Cafe, Coffee Shops are popular in the South East London.
2. As for restaurants, the Italian Restaurants are very popular in the South East London area. Especially in Southwark and Lambeth areas.
3. With the Lewisham area being the most condensed area of Africans in the South East Area, it is surprising to see how in the top 10 venues, you can barely see restaurants in the top 5 venues.
4. Although, the Clusters have variations, a very visible presence is the predominance of pubs.

CONCLUSION AND DISCUSSION

It is very important to note that Clusters 2 and 3 are the most viable clusters to create a brand African Restaurant. Their proximity to other amenities and accessibility to station are paramount. These 2 clusters do not have top restaurants that could rival their standards if they are created. And the proximity to resources needed is paramount as Lewisham and Lambeth are not far out from Peckham (under Southwark).

In conclusion, this project would have had better results if there were more data in terms of crime data within the area, traffic access and allowance of more venues exploration with the Foursquare (limited venues for free calls).

Also, getting the ratings and feedbacks of the current restaurants within the clusters would have helped in providing more insight into the best location.