The Battle of Neighborhoods

Brexit: A Challenge or Opportunity for Catering Industry - Data Analysis of London Restaurant

Introduction – Business Problem

- In 2020, Britain's reckoning with Brexit will leave a significant impact on the hospitality and food industry.

- Take its opportunity to explore expansion after the Brexit

- London is the political, economic, business and cultural art centre

- Diverse and high demand for catering business

- Success Criteria:

- Recommend of borough/neighborhood choice and nearest suppliers of ingredients for organic restaurant

Data Description - 1

- Data Sources (1) - Wikipedia

- The London area consists 32 Boroughs and the "City of London

- The dataset exists for free on the web

- Link to the dataset is https://en.wikipedia.org/wiki/List_of_areas_of_London

	Location	Borough	Post-town	Postcode	Dial-code	OSGridRef
0	Abbey Wood	Bexley, Greenwich	LONDON	SE2	020	TQ465785
1	Acton	Ealing, Hammersmith and Fulham	LONDON	W3, W4	020	TQ205805
2	Addington	Croydon	CROYDON	CR0	020	TQ375645
3	Addiscombe	Croydon	CROYDON	CR0	020	TQ345665
4	Albany Park	Bexley	BEXLEY, SIDCUP	DA5, DA14	020	TQ478728

Data massage and Six Assumptions on Wikipedia Data

We will have multi-rows and assigned the same values from the other columns when the postcode is found more than one

→

We will use the 'Location', 'Borough', 'Postcode', 'Post-town' fields and extracted into a new data frame

→

We assume that the restaurant open for the Boroughs with London Post-town. Therefore, all the non-post-town are dropped

→

Only the South East areas of London will be used for our analysis because of its more diverse outlook, proximity to afro-caribbean markets and accessible facilities. It is given that the South East areas has postcodes with SE.

→

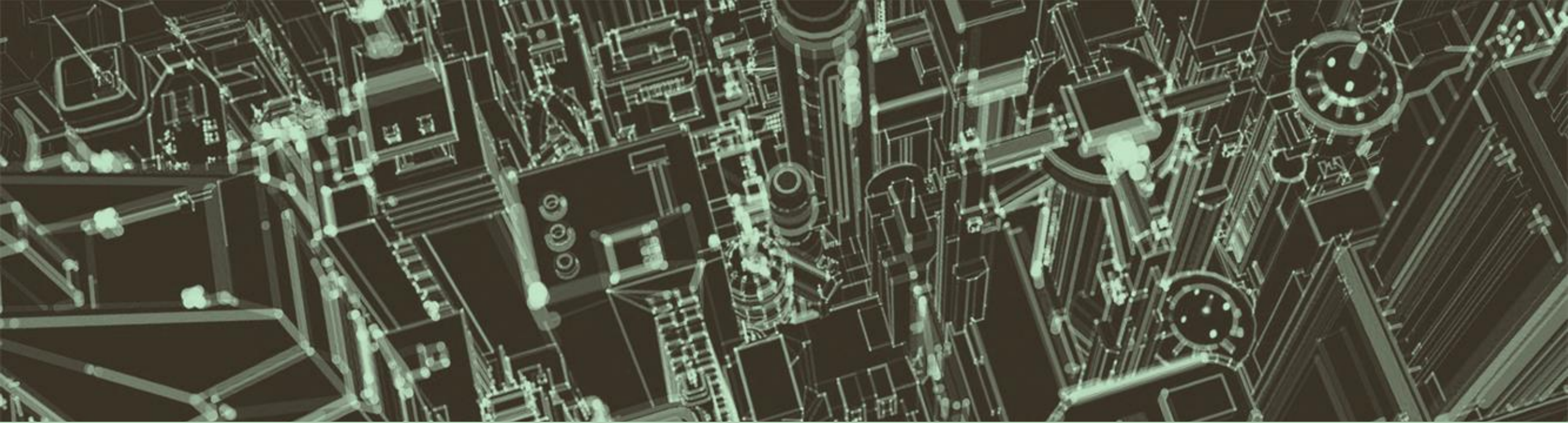We will check the demography of London, mapping the proportion of races by London borough

→

We got a significantly founding from above result, Lewisham, Southward, Lambeth, Hackney and Croydon have a relatively high results in Black.

Data Description - 2

- Data Sources (2) - Foursquare API

- Link for the Foursquare API is https://developer.foursquare.com/

```
{'meta': {'code': 200, 'requestId': '5e5b6205542890001b0c25aa'},
 'response': {'suggestedFilters': {'header': 'Tap to show:',
   'filters': [{'name': 'Open now', 'key': 'openNow'}]},
  'headerLocation': 'Lewisham Central',
  'headerFullLocation': 'Lewisham Central, London',
  'headerLocationGranularity': 'neighborhood',
  'totalResults': 178,
  'suggestedBounds': {'ne': {'lat': 51.47996001800005,
    'lng': 0.021296961190459426},
   'sw': {'lat': 51.44395998200002, 'lng': -0.03637696119035749}},
  'groups': [{'type': 'Recommended Places',
    'name': 'recommended',
    'items': [{'reasons': {'count': 0,
       'items': [{'summary': 'This spot is popular',
         'type': 'general',
         'reasonName': 'globalInteractionReason'}]},
     'venue': {'id': '535823bc498ec8d8da9aad5f',
      'name': 'Street Feast Model Market',
      'location': {'address': '196 Lewisham High St',
       'crossStreet': 'entrance at Molesworth St',
```

Methodology

Exploratory Data Analysis

Single Neighborhood

	name	categories	lat	lng
0	Street Feast Model Market	Street Food Gathering	51.460209	-0.012199
1	Maggie's Kitchen	Café	51.465380	-0.011213
2	Levante restaurant	Restaurant	51.462072	-0.009491
3	Corte	Coffee Shop	51.459776	-0.011554
4	Manor House Gardens	Park	51.456686	0.004684
5	Dirty South	Pub	51.458846	-0.002666
6	Levante Pide Restaurant	Turkish Restaurant	51.459848	-0.011476

	Count
Pub	14
Café	9
Gastropub	7
Park	5
Garden	4

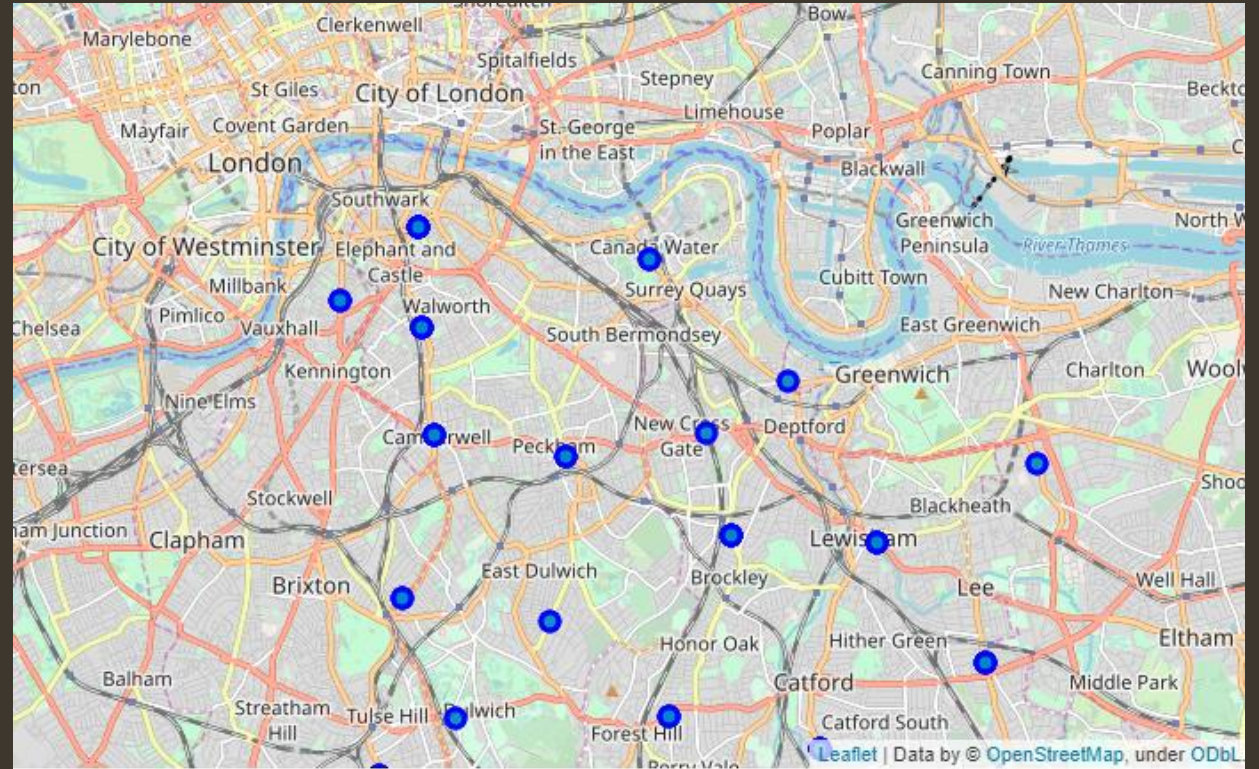## Multiple Neighbourhoods

	Neighbourhood Latitude	Neighbourhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Neighbourhood						
Bankside	200	200	200	200	200	200
Bellingham	134	134	134	134	134	134
Bermondsey	200	200	200	200	200	200
Blackheath	178	178	178	178	178	178
Brixton	200	200	200	200	200	200
Brockley	200	200	200	200	200	200
Camberwell	200	200	200	200	200	200
Catford	134	134	134	134	134	134

	Count
Pub	904
Coffee Shop	604
Café	546
Park	412
Grocery Store	292

Clustering

- The k-means to cluster the neighborhood into 5 clusters

- The top 10 venues can be labelled

- Markers are made in South East London area

Clustering – Example 10 labels & new Dataframe

```
----Bankside----
            venue  freq
0      Coffee Shop  0.10
1              Pub  0.07
2            Hotel  0.07
3  Italian Restaurant  0.05
4          Theater  0.05
5   Scenic Lookout  0.03
6      Cocktail Bar  0.03
7       Art Museum  0.03
8  Seafood Restaurant  0.03
9      Burger Joint  0.02


----Bellingham----
            venue  freq
0     Grocery Store  0.12
1              Park  0.09
2       Supermarket  0.07
3             Café  0.06
4  Fast Food Restaurant  0.04
5  Italian Restaurant  0.04
6              Pub  0.04
7      Coffee Shop  0.04
8      Train Station  0.03
9       Gas Station  0.03
```

```
----Bermondsey----
            venue  freq
0      Coffee Shop  0.10
1              Pub  0.07
2            Hotel  0.07
3  Italian Restaurant  0.05
4          Theater  0.05
5   Scenic Lookout  0.03
6      Cocktail Bar  0.03
7       Art Museum  0.03
8  Seafood Restaurant  0.03
9      Burger Joint  0.02


----Blackheath----
            venue  freq
0              Pub  0.15
1      Coffee Shop  0.06
2             Park  0.06
3     Grocery Store  0.06
4             Café  0.04
5     Clothing Store  0.04
6       Supermarket  0.03
7           Bakery  0.03
8  Italian Restaurant  0.03
9           Garden  0.03
```
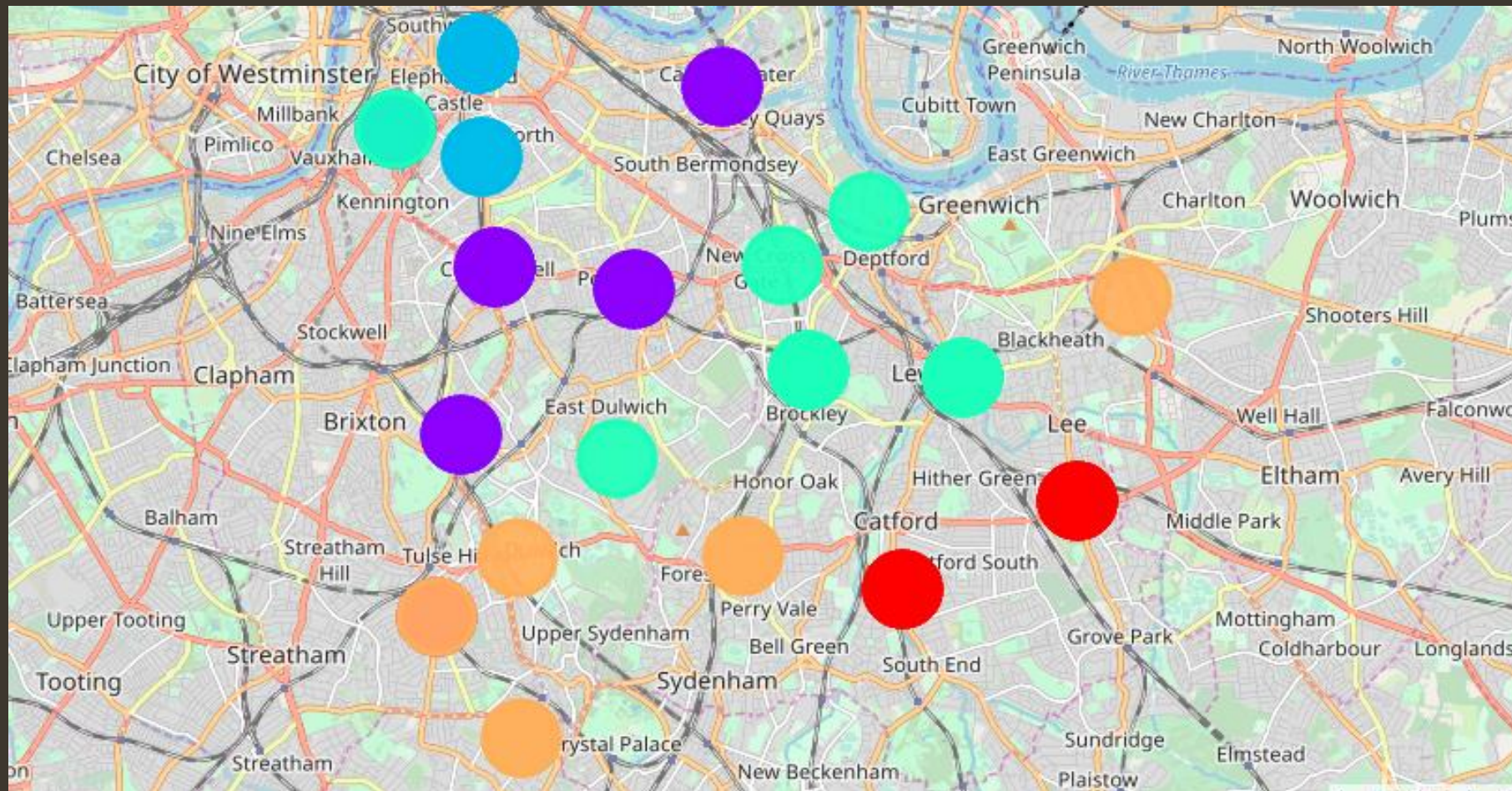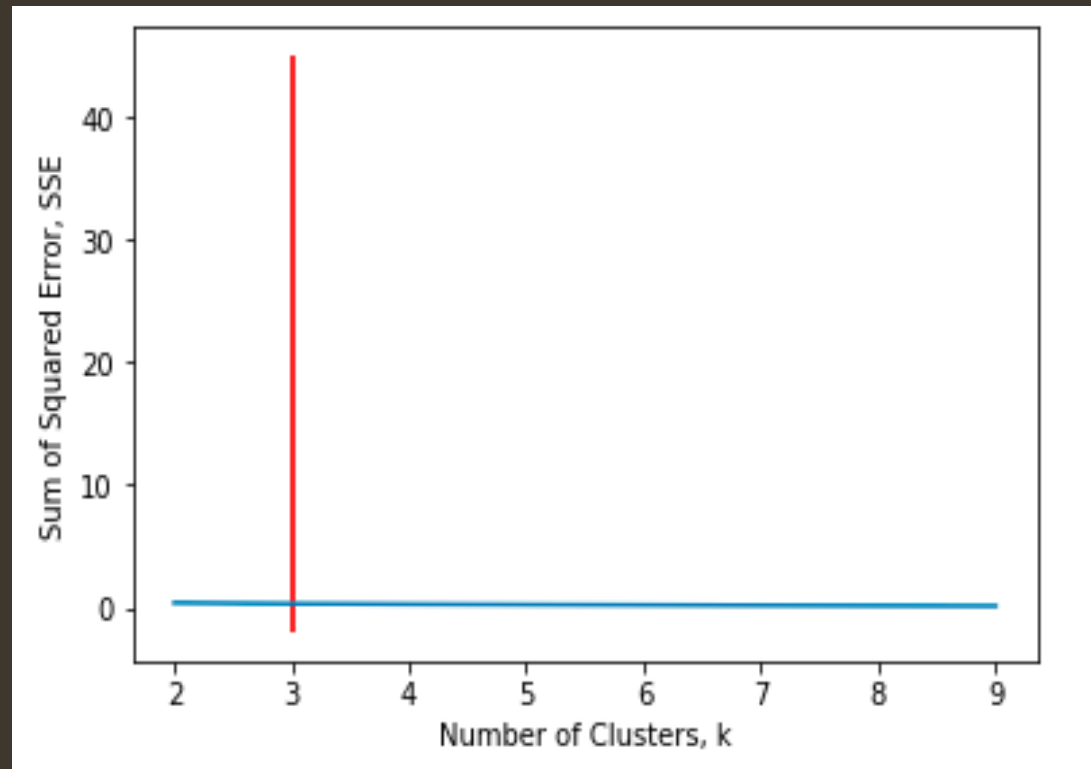
	Location	Borough	Postcode	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue
0	Tulse Hill	Lambeth	SE24	51.45529	-0.09928	1	Coffee Shop	Pub	Café	Grocery Store	Bakery	Park
1	Tulse Hill	Lambeth	SE27	51.43407	-0.10375	1	Coffee Shop	Pub	Café	Grocery Store	Bakery	Park
2	Upper Norwood	Croydon	SE19	51.41990	-0.08808	4	Pub	Café	Park	Coffee Shop	Italian Restaurant	Groc Store
3	Walworth	Southwark	SE17	51.48764	-0.09542	3	Pub	Café	Coffee Shop	Park	Pizza Place	Italia Rest

Visualizing Cluster in Map

Elbow method Result

- The cluster is selected with the minimum SSE.

- 500 iteration are tried

- The optimal k is 3.

Silhouette coefficient Results

- Five clusters are suggested

- 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
17	Lewisham	0	Grocery Store	Pub	Park	Café	Supermarket	Gym / Fitness Center	Italian Restaurant	Indian Restaurant	Train Station	Coffee Shop
24	Lewisham	0	Grocery Store	Pub	Park	Café	Supermarket	Gym / Fitness Center	Italian Restaurant	Indian Restaurant	Train Station	Coffee Shop
31	Lewisham	0	Grocery Store	Park	Supermarket	Café	Coffee Shop	Pub	Fast Food Restaurant	Italian Restaurant	Train Station	Gas Station
37	Lewisham	0	Grocery Store	Park	Supermarket	Café	Coffee Shop	Pub	Fast Food Restaurant	Italian Restaurant	Train Station	Gas Station

- 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
0	Lambeth	1	Coffee Shop	Pub	Café	Grocery Store	Bakery	Park	Pizza Place	Brewery	Gym / Fitness Center	Tapas Restaurant
1	Lambeth	1	Coffee Shop	Pub	Café	Grocery Store	Bakery	Park	Pizza Place	Brewery	Gym / Fitness Center	Tapas Restaurant
6	Southwark	1	Café	Coffee Shop	Pub	Park	Cocktail Bar	Italian Restaurant	Middle Eastern Restaurant	Brewery	Grocery Store	Pizza Place
18	Lambeth	1	Coffee Shop	Café	Pub	Pizza Place	Brewery	Market	Bakery	Burger Joint	Restaurant	Cocktail Bar
29	Southwark	1	Pub	Café	Pizza Place	Coffee Shop	Park	Bar	Indie Movie Theater	Cocktail Bar	Art Gallery	Italian Restaurant

Silhouette coefficient Results

- Cont'

- Cluster 3

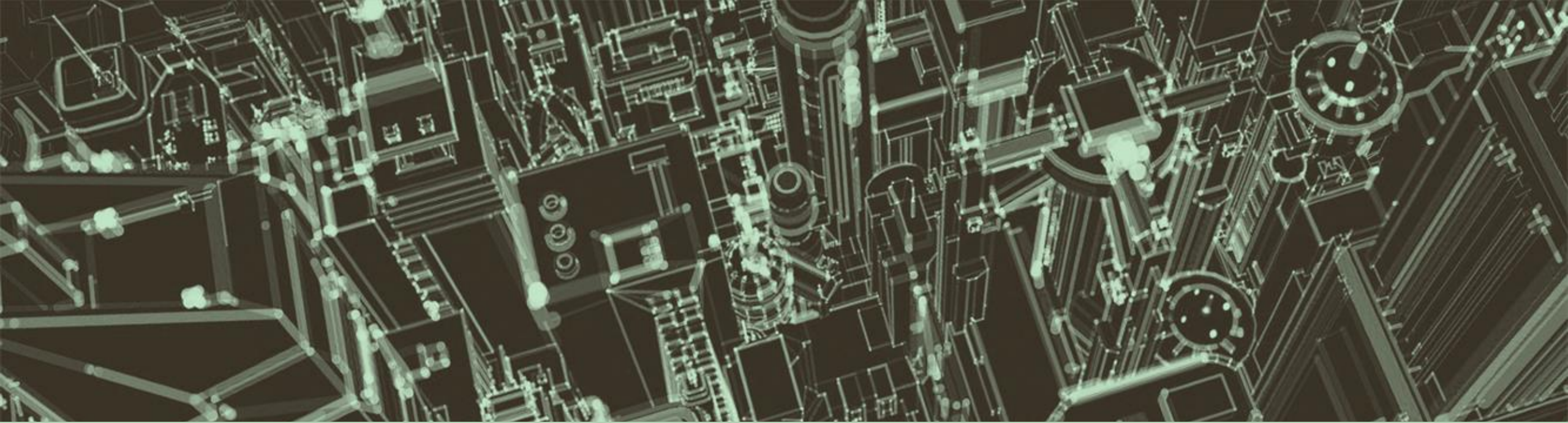|	Borough	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
10	Southwark	2	Pub	Café	Coffee Shop	Hotel	Theater	Park	Italian Restaurant	Street Food Gathering	Pizza Place	Art Gallery
11	Southwark	2	Pub	Café	Coffee Shop	Hotel	Theater	Park	Italian Restaurant	Street Food Gathering	Pizza Place	Art Gallery
12	Southwark	2	Pub	Café	Coffee Shop	Hotel	Theater	Park	Italian Restaurant	Street Food Gathering	Pizza Place	Art Gallery
13	Southwark	2	Coffee Shop	Pub	Hotel	Italian Restaurant	Theater	Scenic Lookout	Cocktail Bar	Seafood Restaurant	Art Museum	Pizza Place

- 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
3	Southwark	3	Pub	Café	Coffee Shop	Park	Pizza Place	Italian Restaurant	Art Gallery	Theater	Cricket Ground	Brewery
5	Lewisham	3	Pub	Coffee Shop	Café	Park	Bar	Gastropub	Indian Restaurant	Cocktail Bar	Italian Restaurant	Hotel
7	Lewisham	3	Pub	Coffee Shop	Café	Bar	Park	Gastropub	Pizza Place	Cocktail Bar	Italian Restaurant	Indie Movie Theater
9	Southwark	3	Pub	Café	Pizza Place	Italian Restaurant	Coffee Shop	Park	Gastropub	Garden Center	Cocktail Bar	Mediterranean Restaurant

Silhouette coefficient Results

- Cont'

- 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
2	Croydon	4	Pub	Café	Park	Coffee Shop	Italian Restaurant	Grocery Store	Train Station	Gastropub	Bakery	Breakfast Spot
4	Lambeth	4	Pub	Coffee Shop	Grocery Store	Café	Park	Bakery	Pizza Place	Gym / Fitness Center	Pharmacy	Indian Restaurant
8	Southwark	4	Pub	Café	Grocery Store	Coffee Shop	Bakery	Park	Brewery	Pizza Place	Farmers Market	Italian Restaurant
14	Lewisham	4	Pub	Coffee Shop	Grocery Store	Café	Park	Supermarket	Gym / Fitness Center	Pizza Place	Japanese Restaurant	Italian Restaurant

Results

Result - Details

- The favorite restaurant in South East London is pubs, café and coffee shop.

- There is a "blue market" for Africans in Lewisham area, South East Area.

- Apart from that, Italian restaurant are famous in South East London Area such as Lambeth area and Southward

Discussion

- Clusters have variation, the most dominance around London is pubs

- In Cluster 2 and 3, it discovered that there is a 'golden nutrient' to create African restaurant.

- Attributed to the proximity to other amenities and accessibility to stations

- The data set and neighborhood can be further expanded

Conclusion

- Analysis is subjected to the limited of data

- Accuracy of the models has room for improvement

- Re-run the program with updated information

- More Ratings and feedbacks, traffic access and allowance obtained from the venues exploration with Foursquare