ONLINE-FOOD-ORDERING

IBM Data Science Capstone Project

MASOUD AYOUBI-NAJAFABADI

Introduction

- **Business Problem:**
- Used the Foursquare API and London data to cluster neighborhoods and determine the best location for the delivery workers of an online food ordering company.

Introduction

- **Background:**
- Online food ordering is a system to order the restaurants' food through the website or mobile app. Based on the type of this system, a customer can choose a restaurant, scan the menu items, select the desired items and pay it.
- What is seriously important for the delivery companies is that they should deliver the food as soon as possible, in order to reduce the delivery time, increase customer's satisfaction, and get more orders.
- It is noticeably vital for these companies to arrange their delivery workers to save the time.
- The idea behind this project is to cluster the neighborhoods of London to ensure that the workers have to drive less distance in each area.
- This project is useful for all the online food delivery companies who would like to increase their performance and profit.

Data

- To answer the business problem, the following data are extracted from the following sources:
  - Population of Each Neighborhood

    (obtained using **London Borough Profiles Database)**

  - Coordinates of the Neighborhoods of London

    (obtained by **parsing the Wikipedia web page: List_of_London_boroughs)**

  - Number of Restaurants in Each Neighborhood

    (obtained using **Foursquare API)**

Data

- To answer the business problem, the following data are extracted from the following sources:
  - Population of Each Neighborhood

    (obtained using **London Borough Profiles Database)**

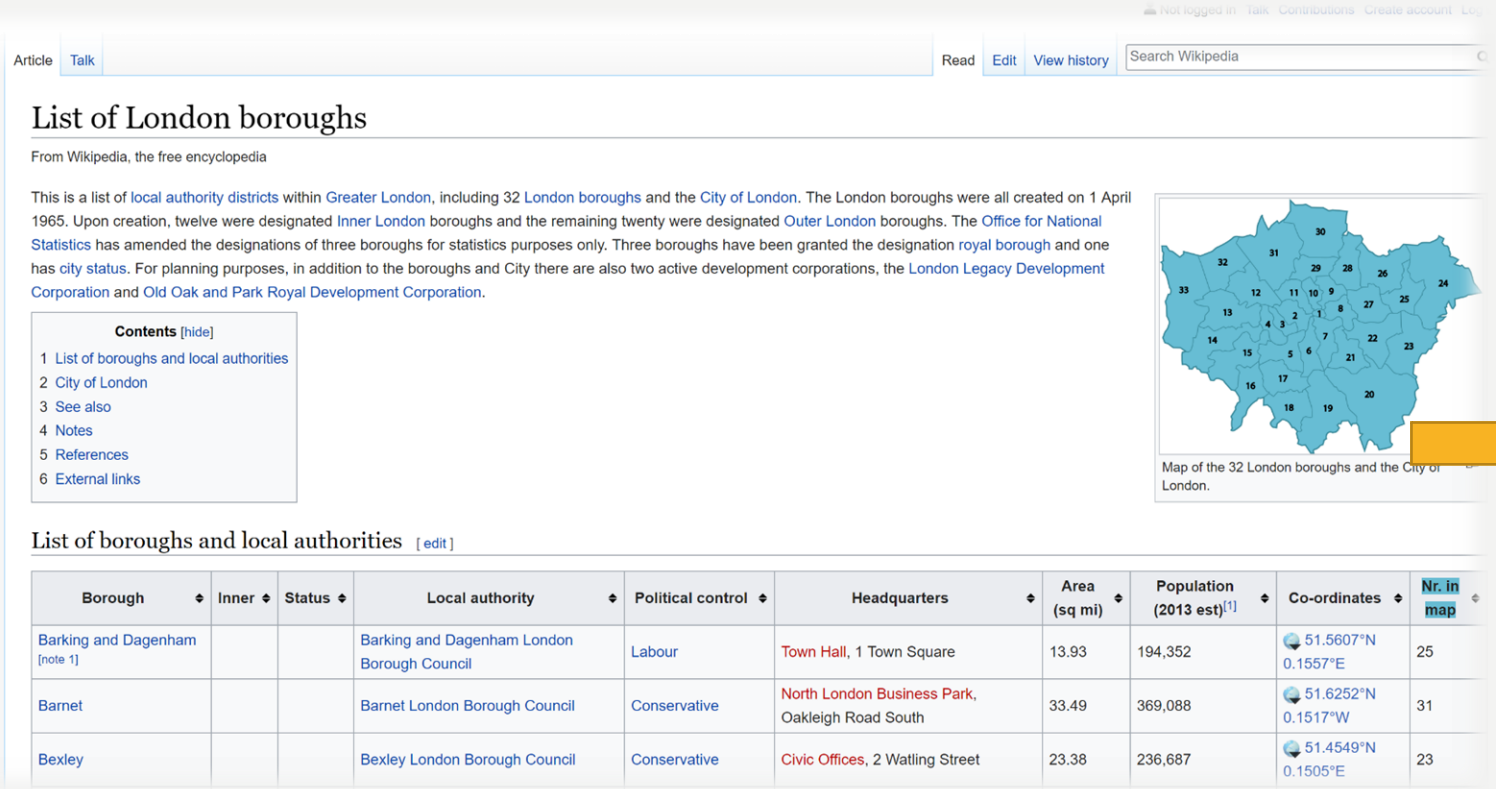
	Code	Area_name	Inner/_Outer_London	GLA_Population_Estimate_2017	GLA_Household_Estimate_2017	Inland_Area_(Hectares)	Population_density_(per_hectare)_2017	A
0	E09000001	City of London	Inner London	8800	5326	290	30.3	
1	E09000002	Barking and Dagenham	Outer London	209000	78188	3,611	57.9	
2	E09000003	Barnet	Outer London	389600	151423	8,675	44.9	
3	E09000004	Bexley	Outer London	244300	97736	6,058	40.3	
4	E09000005	Brent	Outer London	332100	121048	4,323	76.8	
5	E09000006	Bromley	Outer London	327900	140602	15,013	21.8	
6	E09000007	Camden	Inner London	242500	107654	2,179	111.3	
7	E09000008	Croydon	Outer London	386500	159010	8,650	44.7	
8	E09000009	Ealing	Outer London	351600	132663	5,554	63.3	
9	E09000010	Enfield	Outer London	333000	130328	8,083	41.2	
10	E09000011	Greenwich	Outer London	280100	113964	4,733	59.2	
11	E09000012	Hackney	Inner London	274300	115417	1,905	144	

	Borough	Population
0	City of London	8800
1	Barking and Dagenham	209000
2	Barnet	389600
3	Bexley	244300
4	Brent	332100
5	Bromley	327900
6	Camden	242500
7	Croydon	386500
8	Ealing	351600
9	Enfield	333000
10	Greenwich	280100
11	Hackney	274300
12	Hammersmith and Fulham	185300

Data

- To answer the business problem, the following data are extracted from the following sources:
  - Coordinates of the Neighborhoods of London

    (obtained by **parsing the Wikipedia web page: List_of_London_boroughs)**



Wikipedia
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Wikipedia store

Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact page

Tools
What links here
Related changes
Upload file
Special pages
Permanent link
Page information
Wikidata item
Cite this page

Print/export
Create a book
Download as PDF

Article Talk Read Edit View history Search Wikipedia

List of London boroughs

From Wikipedia, the free encyclopedia

This is a list of local authority districts within Greater London, including 32 London boroughs and the City of London. The London boroughs were all created on 1 April 1965. Upon creation, twelve were designated Inner London boroughs and the remaining twenty were designated Outer London boroughs. The Office for National Statistics has amended the designations of three boroughs for statistics purposes only. Three boroughs have been granted the designation royal borough and one has city status. For planning purposes, in addition to the boroughs and City there are also two active development corporations, the London Legacy Development Corporation and Old Oak and Park Royal Development Corporation.

Contents [hide]
1 List of boroughs and local authorities
2 City of London
3 See also
4 Notes
5 References
6 External links

Map of the 32 London boroughs and the City of London.

List of boroughs and local authorities [edit]

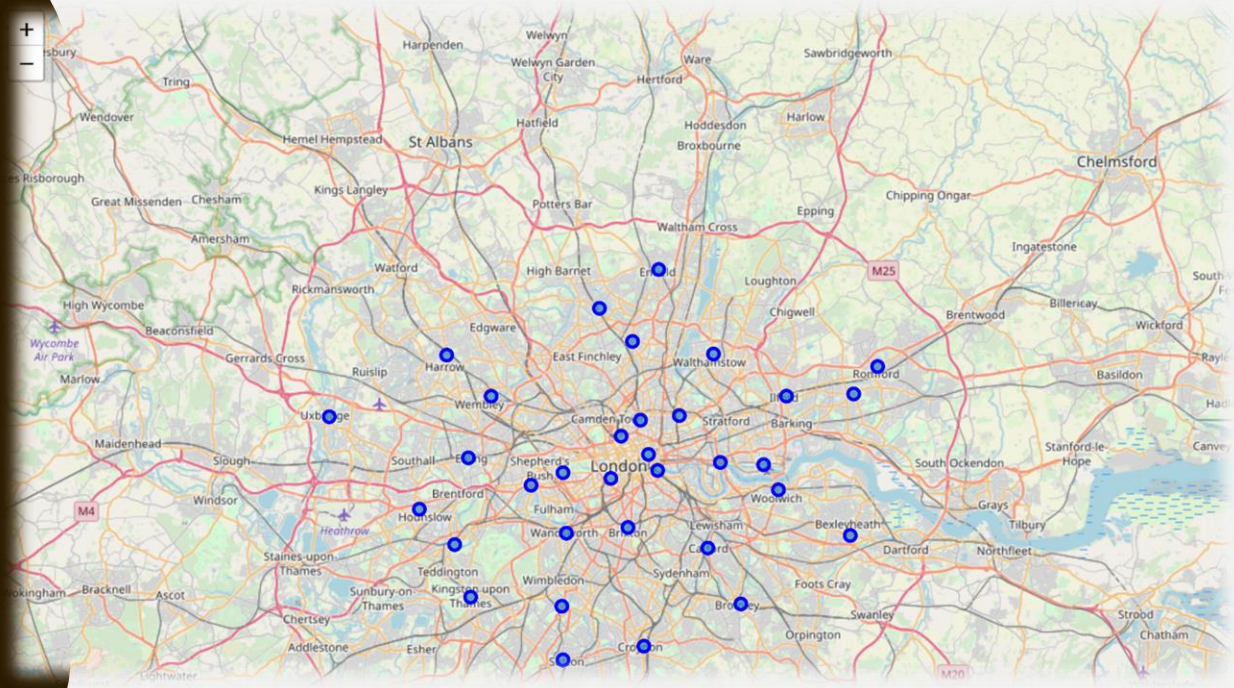
Borough	Inner	Status	Local authority	Political control	Headquarters	Area (sq mi)	Population (2013 est)[1]	Co-ordinates	Nr. in map
Barking and Dagenham [note 1]			Barking and Dagenham London Borough Council	Labour	Town Hall, 1 Town Square	13.93	194,352	51.5607°N 0.1557°E	25
Barnet			Barnet London Borough Council	Conservative	North London Business Park, Oakleigh Road South	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

	Borough	Latitude	Longitude
0	Barking and Dagenham	51.5607	0.1557
1	Barnet	51.6252	-0.1517
2	Bexley	51.4549	0.1505
3	Brent	51.5588	-0.2817
4	Bromley	51.4039	0.0198
5	Camden	51.5290	-0.1255
6	Croydon	51.3714	-0.0977
7	Ealing	51.5130	-0.3089
8	Enfield	51.6538	-0.0799
9	Greenwich	51.4892	0.0648
10	Hackney	51.5450	-0.0553
11	Hammersmith and Fulham	51.4927	-0.2339
12	Haringey	51.6000	-0.1119
13	Harrow	51.5898	-0.3346
14	Havering	51.5812	0.1837
15	Hillingdon	51.5441	-0.4760
16	Hounslow	51.4746	-0.3680
17	Islington	51.5416	-0.1022

Data

- To answer the business problem, the following data are extracted from the following sources:
  - Number of Restaurants in Each Neighborhood

    (obtained using **Foursquare API)**



Neighborhood	Number of venues
Barking and Dagenham	7
Barnet	55
Bexley	40
Brent	74
Bromley	43
Camden	100
City of London	100
Croydon	79
Ealing	100
Enfield	37
Greenwich	45
Hackney	100
Hammersmith and Fulham	100
Haringey	61
Harrow	46
Havering	37
Hillingdon	44
Hounslow	51

Data

- Merging and cleaning data:

Neighborhood	Number of v
Barking and Dagenham	
Barnet	
Bexley	
Brent	
Bromley	
Camden	
City of London	
Croydon	
Ealing	
Enfield	
Greenwich	
Hackney	
Hammersmith and Fulham	
Haringey	
Harrow	
Havering	37
Hillingdon	44
Hounslow	51

	Borough	Latit	
0	Barking and Dagenham	51.5	
1	Barnet	51.6	
2	Bexley	51.4	
3	Brent	51.5	
4	Bromley	51.4	
5	Camden	51.5	
6	Croydon	51.3	
7	Ealing	51.5	
8	Enfield	51.6	
9	Greenwich	51.4892	0.0648
10	Hackney	51.5450	-0.0553
11	Hammersmith and Fulham	51.4927	-0.2339
12	Haringey	51.6000	-0.1119
13	Harrow	51.5898	-0.3346
14	Havering	51.5812	0.1837
15	Hillingdon	51.5441	-0.4760
16	Hounslow	51.4746	-0.3680
17	Islington	51.5416	-0.1022

	Borough	Population
0	City of London	8800
1	Barking and Dagenham	209000
2	Barnet	389600
3	Bexley	244300
4	Brent	332100
5	Bromley	327900
6	Camden	242500
7	Croydon	386500
8	Ealing	351600
9	Enfield	333000
10	Greenwich	280100
11	Hackney	274300
12	Hammersmith and Fulham	185300

	Neighborhood	Population	Latitude	Longitude	Cluster Labels	Number of venues
0	City of London	8800	51.5155	-0.0922	5	100
1	Barking and Dagenham	209000	51.5607	0.1557	0	7
2	Barnet	389600	51.6252	-0.1517	3	55
3	Bexley	244300	51.4549	0.1505	4	40
4	Brent	332100	51.5588	-0.2817	2	74
5	Bromley	327900	51.4039	0.0198	1	43
6	Camden	242500	51.5290	-0.1255	5	100
7	Croydon	386500	51.3714	-0.0977	2	79
8	Ealing	351600	51.5130	-0.3089	1	100
9	Enfield	333000	51.6538	-0.0799	2	37
10	Greenwich	280100	51.4892	0.0648	1	45
11	Hackney	274300	51.5450	-0.0553	3	100
12	Hammersmith and Fulham	185300	51.4927	-0.2339	1	100
13	Haringey	278000	51.6000	-0.1119	3	61
14	Harrow	252300	51.5898	-0.3346	2	46
15	Havering	254300	51.5812	0.1837	6	37
16	Hillingdon	301000	51.5441	-0.4760	4	44
17	Hounslow	274200	51.4746	-0.3680	7	51
18	Islington	231200	51.5416	-0.1022	5	100

Methodology

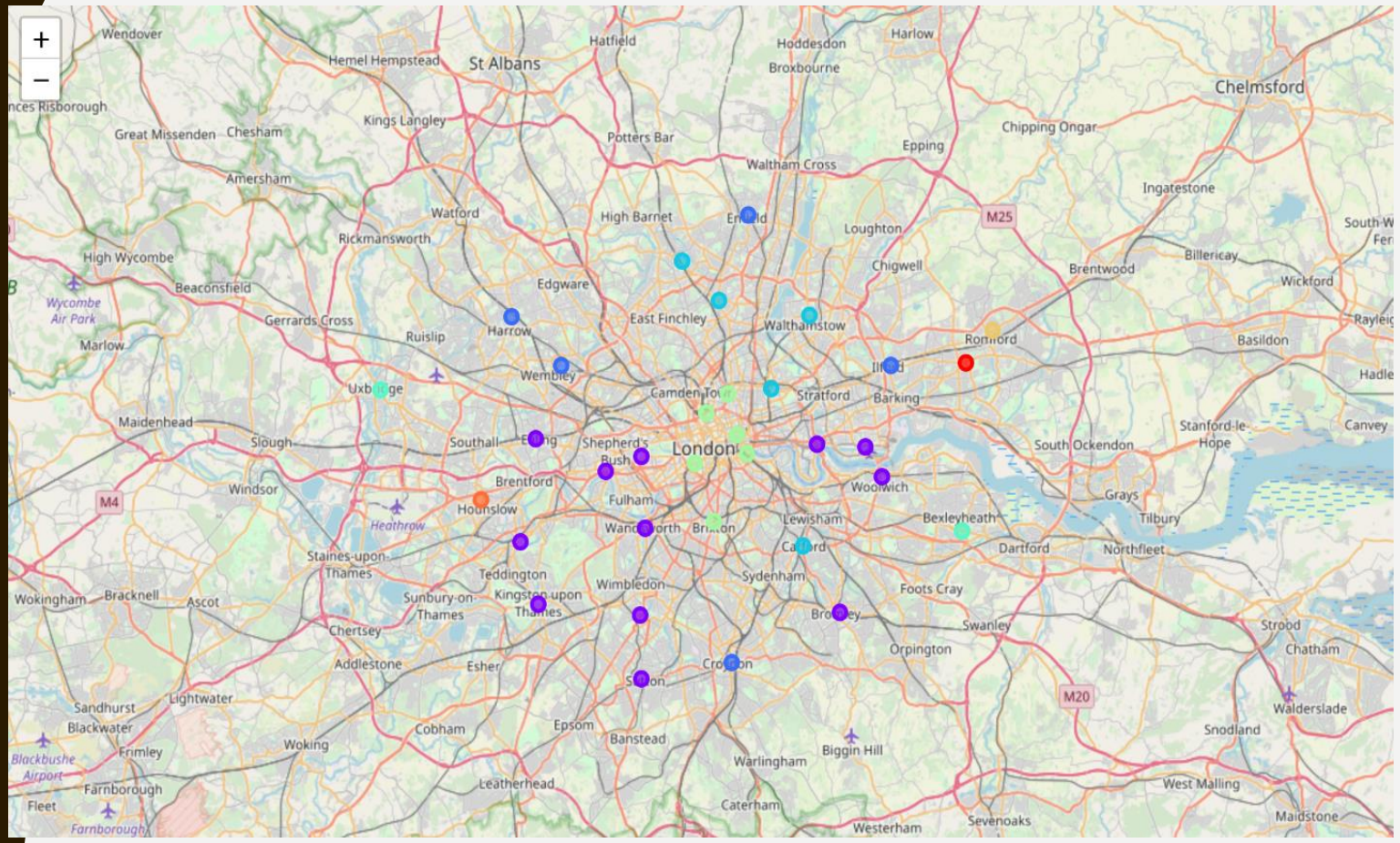
- The scope of this project is to determine the areas of London with most food category venues (restaurant, fast food, café ,...) density, particularly those in crowded areas with high population.
  - In first step we have collected the required data: location and type (category) of food category venues (according to Foursquare categorization) and population of each area.
  - Second step in our analysis will be calculation and exploration of 'venue density' across different areas of London.
  - In the third step, we will focus on most promising areas and within those create clusters of locations. We will take into consideration locations in radius of 2000 meters. We will present map of all such locations but also create clusters (using k-means clustering) of those locations to identify neighborhoods. Then, the scatter plot of population and number of each venue for each cluster is presented.

Results and Discussion

- The analysis indicates that although there are 33 boroughs in London having different food venues and population, some of these areas are more crowded in comparison to the other ones.
- The density of the food venues in these areas are completely different.
- This analysis demonstrates the importance of this survey for the stakeholders.
- Also, it clarified that this diversity among the areas can widely influence the density and location of delivery workers in the city.

Results and Discussion

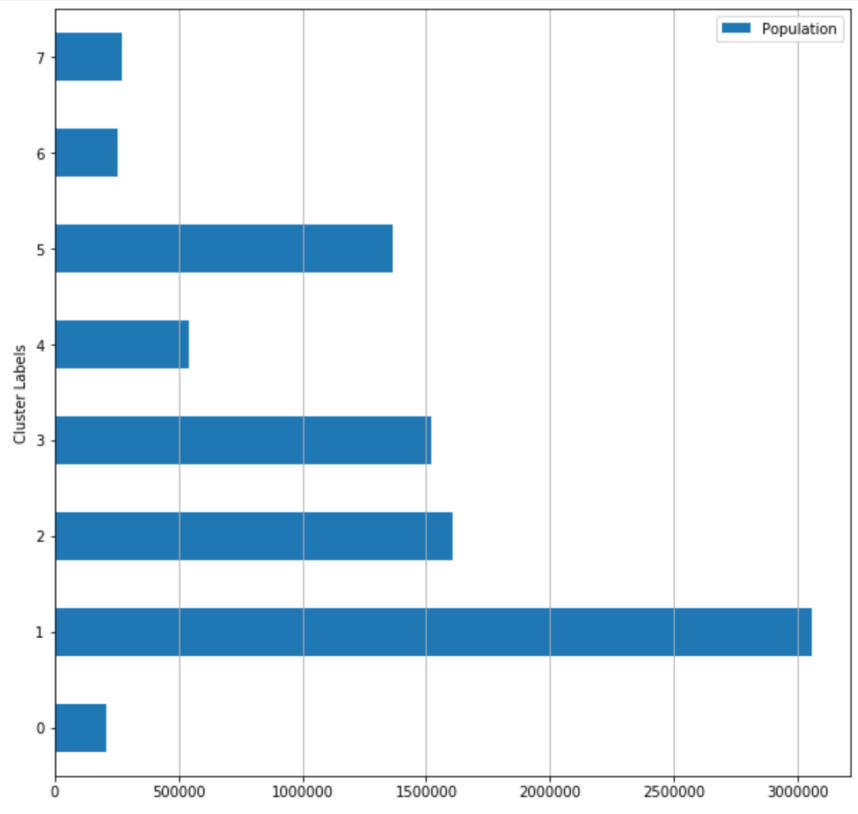
- The k-means method is used to clustering the neighborhoods into 8 clusters. The areas were illustrated in the map categorizing in the groups with various colors.



Cluster Labels	Neighborhood
0	Barking and Dagenham
1	Bromley, Ealing, Greenwich, Hammersmith and Fu...
2	Brent, Croydon, Enfield, Harrow, Redbridge
3	Barnet, Hackney, Haringey, Lewisham, Waltham F...
4	Bexley, Hillingdon
5	City of London, Camden, Islington, Lambeth, So...
6	Havering
7	Hounslow

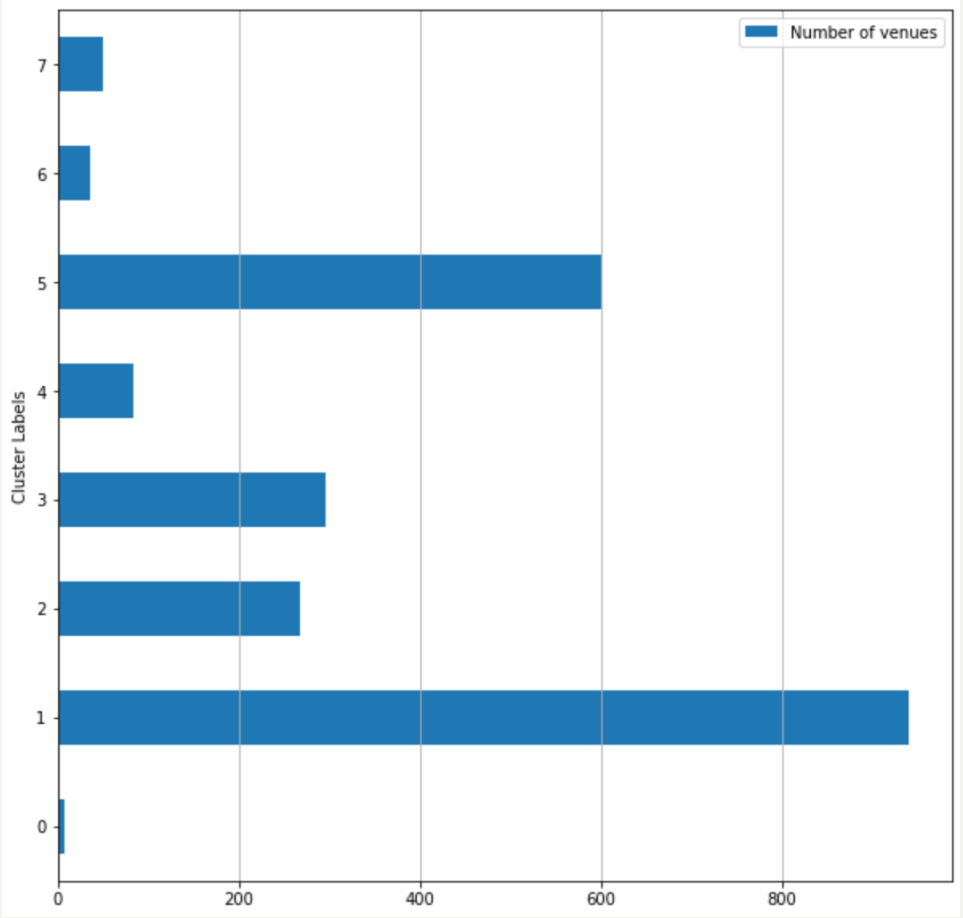
Results and Discussion

- The crowded areas are clusters labeled 3 and 2, while the clusters labeled 4, 6 and 7 have the least population. If we consider only the population of each area, the results show that the crowded areas could be potentially good places to arrange the most delivery workers, but it is not the only case we should consider.



Results and Discussion

- Generally, we expect that those areas with more venues need more online services. As illustrated in the "Number of Venues" figure, cluster 0 and cluster 3 have the most food venues, while clusters 4, 6 and 7 have the least ones.



Results and Discussion

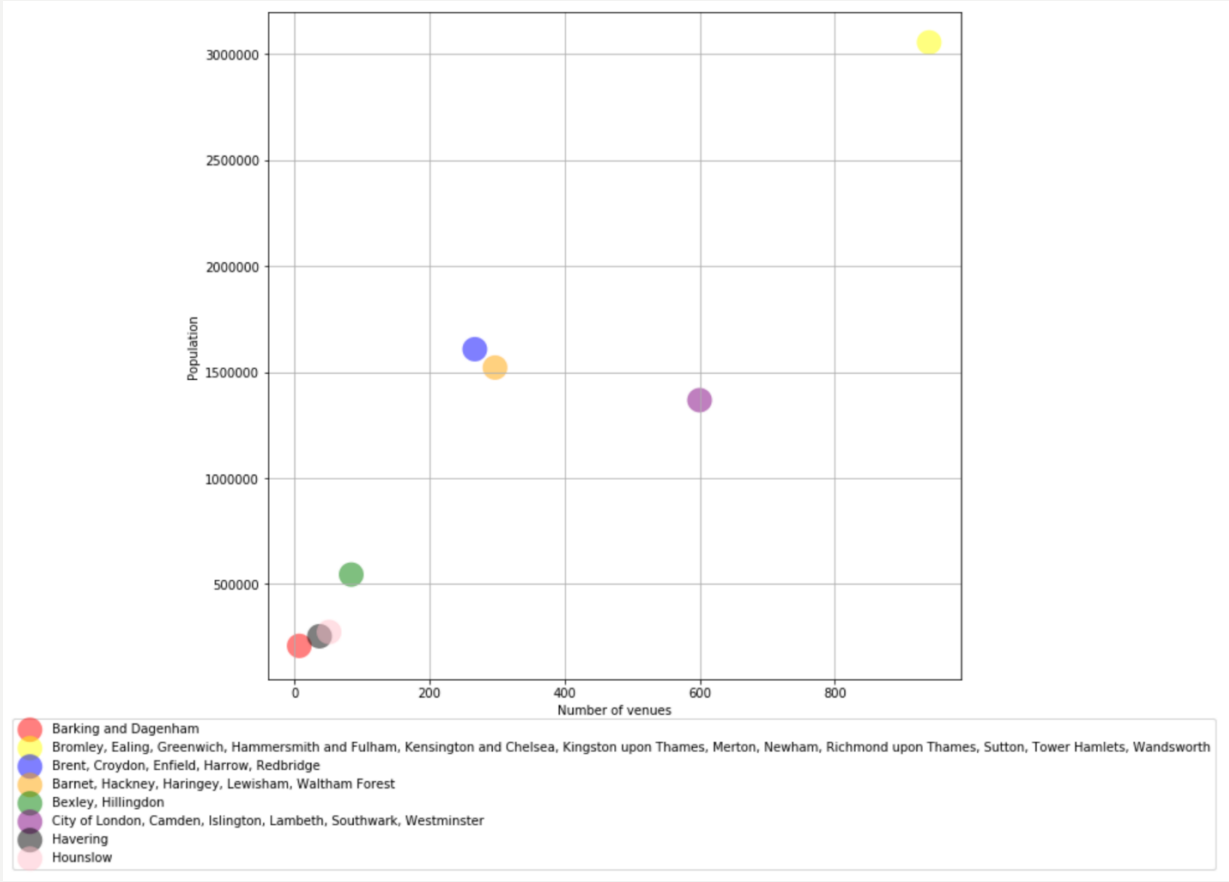
- As we can see, if we consider the two variables ("population" and "number of food venues") one by one, it could give us a good estimation, but not a perfect result. Because, the high number of food venues in an area cannot guarantee that the more people would like use online services, number of food venues itself could not be a variable with a hundred percent accuracy.

- If an area has more population, it does not ensure that always more restaurants and fast foods are there.

- Therefore, to estimate the tendency of the customers for ordering online food, it is necessary to evaluate both variables in each area together.

Results and Discussion

- The final results provided in the scatter plot reveals the relation between "population" and "number of food venues" in each neighborhood. As we can see, the best place to integrate the workers is cluster 3. After that, we recommend that the company focuses on cluster 2, cluster 0 and cluster 5.



Conclusion

- The scope of this project was to identify London areas with high number of food venues in order to assist online food ordering stakeholders in narrowing down the search for optimal location for arranging delivery workers.

- By calculating restaurant density distribution from Foursquare data we have first identified general boroughs that justify further analysis, then generated extensive collection of locations considering the population of each area.

- Clustering of those locations was then performed in order to create major zones of interest.