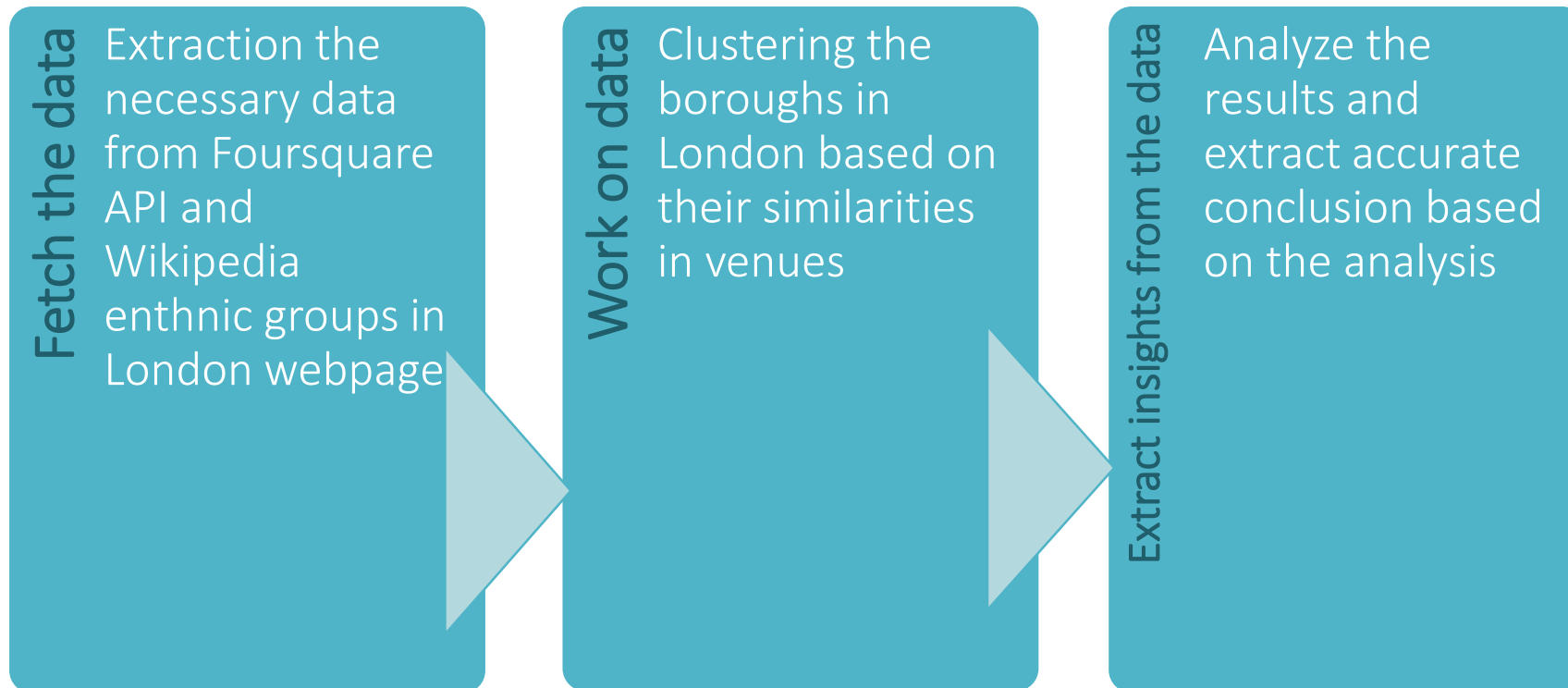# FINDING THE OPTIMAL LOCATION FOR A BUSINESS

Fotina Nataliya

Coursera_Applied Data Science Capstone course

Main steps for location determination

This project initiated for determination a location for starting a business in London city area.

Fetch the data — Extraction the necessary data from Foursquare API and Wikipedia enthnic groups in London webpage

Work on data — Clustering the boroughs in London based on their similarities in venues

Extract insights from the data — Analyze the results and extract accurate conclusion based on the analysis

Distribution of various Asian ethnic groups in London

The data that was used contained the information about distribution of various Asian ethnicities with respect to their location in London:

Rank	London Borough	Indian Population	Pakistani Population	Bangladeshi Population	Chinese Population	
1	Newham	42,484	30,307	37,262	3,930	1
2	Redbridge	45,660	31,051	16,011	3,000	2
3	Brent	58,017	14,381	1,749	3,250	2
4	Tower Hamlets	6,787	2,442	81,377	8,109	5
5	Harrow	63,051	7,797	1,378	2,829	2
6	Ealing	48,240	14,711	1,786	4,132	3
7	Hounslow	48,161	13,679	2,169	2,405	2
8	Hillingdon	36,795	9,200	2,639	2,889	1
9	Barnet	27,920	5,344	2,215	8,259	2
10	Croydon	24,660	10,865	2,570	3,925	1
11	Waltham Forest	9,134	26,347	4,632	2,579	1
12	Merton	8,106	7,337	2,218	2,818	1
13	Camden	6,083	1,489	12,503	6,493	8
14	Enfield	11,848	2,594	5,389	2,588	1
15	Wandsworth	8,642	9,718	1,493	3,715	9
16	Westminster	7,213	2,328	6,299	5,917	1
17	Greenwich	7,836	2,594	1,645	5,061	1
18	Barking and Dagenham	7,436	8,007	7,701	1,315	5
19	Southwark	5,819	1,823	3,912	8,074	7
20	Kingston Upon Thames	6,325	3,009	892	2,883	1

Foursquare API

Location data was appended to the scraped data in order to access Foursquare API.

Second DataFrame was created with neighborhood names.

	Neighborhood	Latitude	Longitude
0	Newham	51.5255	0.0352
1	Redbridge	51.5901	0.0819
2	Brent	51.5673	-0.2711
3	Tower Hamlets	51.5203	-0.0293
4	Harrow	51.5806	-0.3420

Information about latitude and longitude of each neighbourhood was added to dataframe.

London map

The map was generated consisting of all the neighborhoods

Venues by neighborhood

Neighboring objects by category were found in relation to each borough. Nearby venues was organized in using location data.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Newham	51.5255	0.0352	Delicious Café	51.526417	0.030133	Café
1	Newham	51.5255	0.0352	Tesco Express	51.527187	0.035118	Grocery Store
2	Newham	51.5255	0.0352	Andre Moves	51.524192	0.036145	Home Service
3	Newham	51.5255	0.0352	Deep Blue Sea Fish & Chips	51.525097	0.039410	Fish & Chips Shop
4	Newham	51.5255	0.0352	Ginny's Pie and Mash	51.525705	0.029532	Café

Frequency of occurrence

Nearby venues for each neighborhood obtained with their frequencies of occurrences:

```
----Barnet----
venue    freq
          Café
       Bus Stop
merican Restaurant
Recreation Center
      Public Art
```

K-Means

Next were generated clusters using K-Means:

Cluster Three:

In [134]: `london_merged.loc[london_merged['Cluster Labels'] == 2, london_merged.columns[[0] + list(range(1, london_merged.shape[`

Out[134]:

	London Borough	Indian Population	Pakistani Population	Bangladeshi Population	Chinese Population	Other Asian Population	Total Asian Population	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
15	Westminster	7213	2328	6299	5917	10105	31862	Westminster	Hotel	Coffee Shop	Sandwich Place	Sushi Restaurant	Theater

Cluster Four:

In [135]: `london_merged.loc[london_merged['Cluster Labels'] == 3, london_merged.columns[[0] + list(range(1, london_merged.shape[`

Out[135]:

	London Borough	Indian Population	Pakistani Population	Bangladeshi Population	Chinese Population	Other Asian Population	Total Asian Population	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
13	Enfield	11848	2594	5599	2588	12464	34893	Enfield	Pub	Coffee Shop	Restaurant	Auto Workshop	Tennis Court
10	Waltham Forest	9134	26347	4632	2579	11697	54389	Waltham Forest	Grocery Store	Pub	Coffee Shop	Concert Hall	Vegetarian / Vegan Restaurant
17	Barking and Dagenham	7436	8007	7701	1315	5135	29594	Barking and Dagenham	Lake	Park	Women's Store	Food Court	Department Store
12	Camden	6083	1489	12503	6493	8878	35446	Camden	Gastropub	Bakery	Pizza Place	Coffee Shop	Café
18	Southwark	5819	1623	3912	8074	7764	27192	Southwark	Pub	Building	Café	Skate Park	Park

Visualization

Clusters were plotted on the map for visualization: