

WHERE TO OPEN A PAKISTANI RESTAURANT IN LONDON?

Capstone Project Presented by:

Talha Bin Khalid



CONTENTS

- Business Problem
- Data Collection
- Methodology
- Results
- Conclusion



BUSINESS PROBLEM

- London – a multicultural tourist destination
- Large number of Pakistani Population in London
- What would be the most suitable location to open a Pakistani Restaurant in London?
- The location needs to have enough influx of relevant customers
- Exploring 'Asian Restaurants' in London can give us an indication of which locations would be suitable for a Pakistani restaurant



DATA COLLECTION

- Two types of data sources were used:
 - Location (Neighborhood) data of London
 - Demographic data of London (Proportion of races in different neighborhoods)
- Foursquare venue data was used in the methodology section



LOCATION DATA OF LONDON

- Using beautifulsoup package, the neighborhood data for London was extracted from the Wikipedia page:
 - https://en.wikipedia.org/wiki/List_of_areas_of_London

- Cleaned data:

	Location	Borough	Post town	Dial code	OS grid ref	Postcode
0	Abbey Wood	Bexley, Greenwich	LONDON	020	TQ465785	SE2
1	Acton	Ealing, Hammersmith and Fulham	LONDON	020	TQ205805	W3
1	Acton	Ealing, Hammersmith and Fulham	LONDON	020	TQ205805	W4
2	Addington	Croydon	CROYDON	020	TQ375645	CR0
3	Addiscombe	Croydon	CROYDON	020	TQ345665	CR0
4	Albany Park	Bexley	BEXLEY, SIDCUP	020	TQ478728	DA5
4	Albany Park	Bexley	BEXLEY, SIDCUP	020	TQ478728	DA14
5	Aldborough Hatch	Redbridge	ILFORD	020	TQ455895	IG2
6	Aldgate	City	LONDON	020	TQ334813	EC3
7	Aldwych	Westminster	LONDON	020	TQ307810	WC2



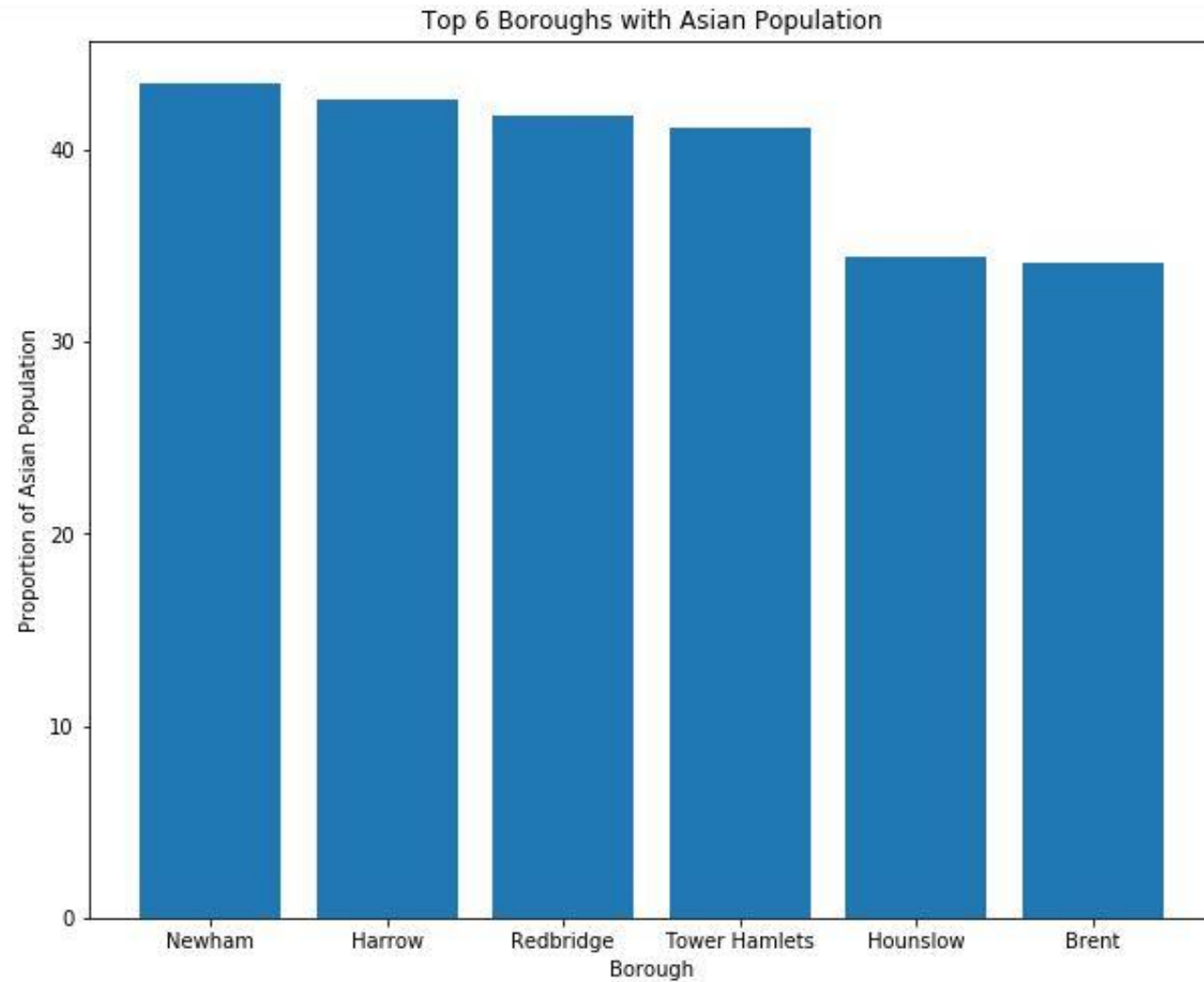
DEMOGRAPHIC DATA OF LONDON

- Using beautifulsoup package, the demographic data for London was extracted from the Wikipedia page:
 - https://en.wikipedia.org/wiki/Demography_of_London
- We chose the top 6 neighborhoods with the highest proportion (more than 30%) of Asians:

	Local authority	White	Mixed	Asian	Black	Other
24	Newham	29	4.5	43.5	19.6	3.5
13	Harrow	42.2	4	42.6	8.2	2.9
25	Redbridge	42.5	4.1	41.8	8.9	2.7
29	Tower Hamlets	45.2	4.1	41.1	7.3	2.3
17	Hounslow	51.4	4.1	34.4	6.6	3.6
3	Brent	36.3	5.1	34.1	18.8	5.8



DEMOGRAPHIC DATA



LOCATION COORDINATES

- Arcgis_geocoder was used to extract the location coordinates for our relevant neighborhoods
- Location coordinates were split into latitude and longitude columns in a dataframe
- Latitude and Longitude data was merged with the original dataframe of neighborhoods



MERGED DATAFRAME

	Location	Borough	Postcode	Latitude	Longitude
0	Beckton	Newham	E6	51.53292	0.054610
1	Beckton	Newham	E16	51.50913	0.015280
2	Beckton	Newham	IG11	51.53312	0.084077
3	Bethnal Green	Tower Hamlets	E2	51.52669	-0.062570
4	Blackwall	Tower Hamlets	E14	51.51122	-0.012640



METHODOLOGY

- Foursquare API was used to explore the nearby venues from our neighborhoods
- Initially 'Blackwall' neighborhood was used to explore the nearby venues and the most common venue
- Coffee Shop came out to be the most common venue for Blackwall
- The same process was repeated using foursquare API for all the neighborhoods to get the nearby venues

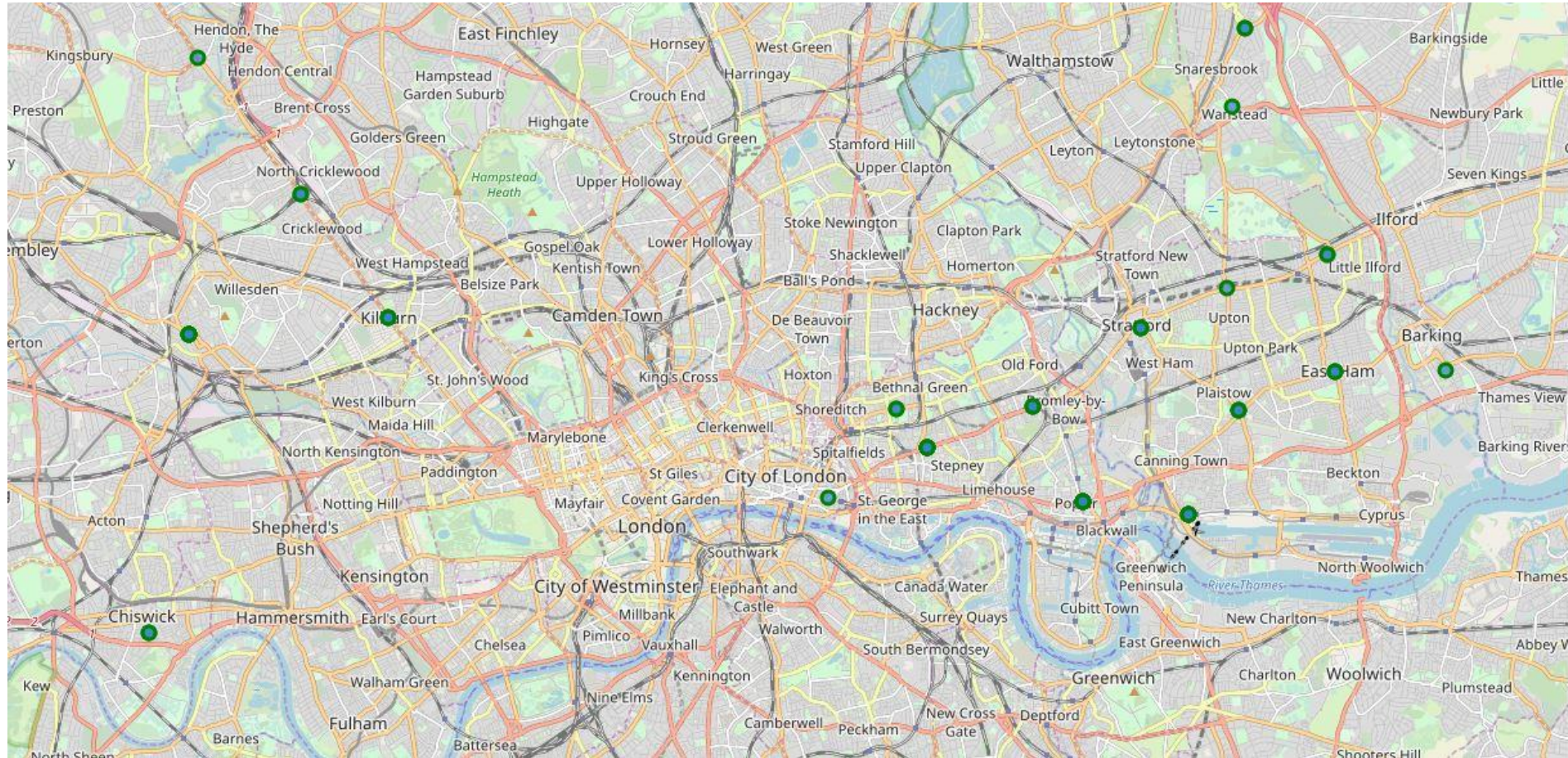


NEARBY VENUES FOR NEIGHBORHOODS

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Beckton	51.53292	0.05461	McDonald's	51.534031	0.053797	Fast Food Restaurant
1	Beckton	51.53292	0.05461	The Miller's Well (Wetherspoon)	51.533406	0.056379	Pub
2	Beckton	51.53292	0.05461	Central Park	51.528808	0.052901	Park
3	Beckton	51.53292	0.05461	Costa Coffee	51.534517	0.053365	Coffee Shop
4	Beckton	51.53292	0.05461	Primark	51.535303	0.052308	Clothing Store



VISUALIZATION



Map of London with marked top 6 Asian neighborhoods



EXPLORATORY DATA ANALYSIS

- The nearby venue categories were extracted
- All the 222 unique venue categories were set into columns
- The neighborhoods were grouped and mean of each venue category was calculated for each neighborhood
- The list of venue categories was observed closely which led to a change of initially defined methodology



EXPLORATORY DATA ANALYSIS

- At the start of this project, it was described that the criteria for finding a suitable location for a Pakistani restaurant will be to analyze the locations having the most number of 'Asian Restaurants'
- But the significant presence of Indian Restaurants in the venue category led to change in the methodology
- Indians and Pakistanis belong to the same ethnicity and have similar cuisines
- It would be better for our analysis to explore 'Indian Restaurants' rather than 'Asian Restaurants'
- So the 'Indian Restaurants' will be explored further in our analysis ahead



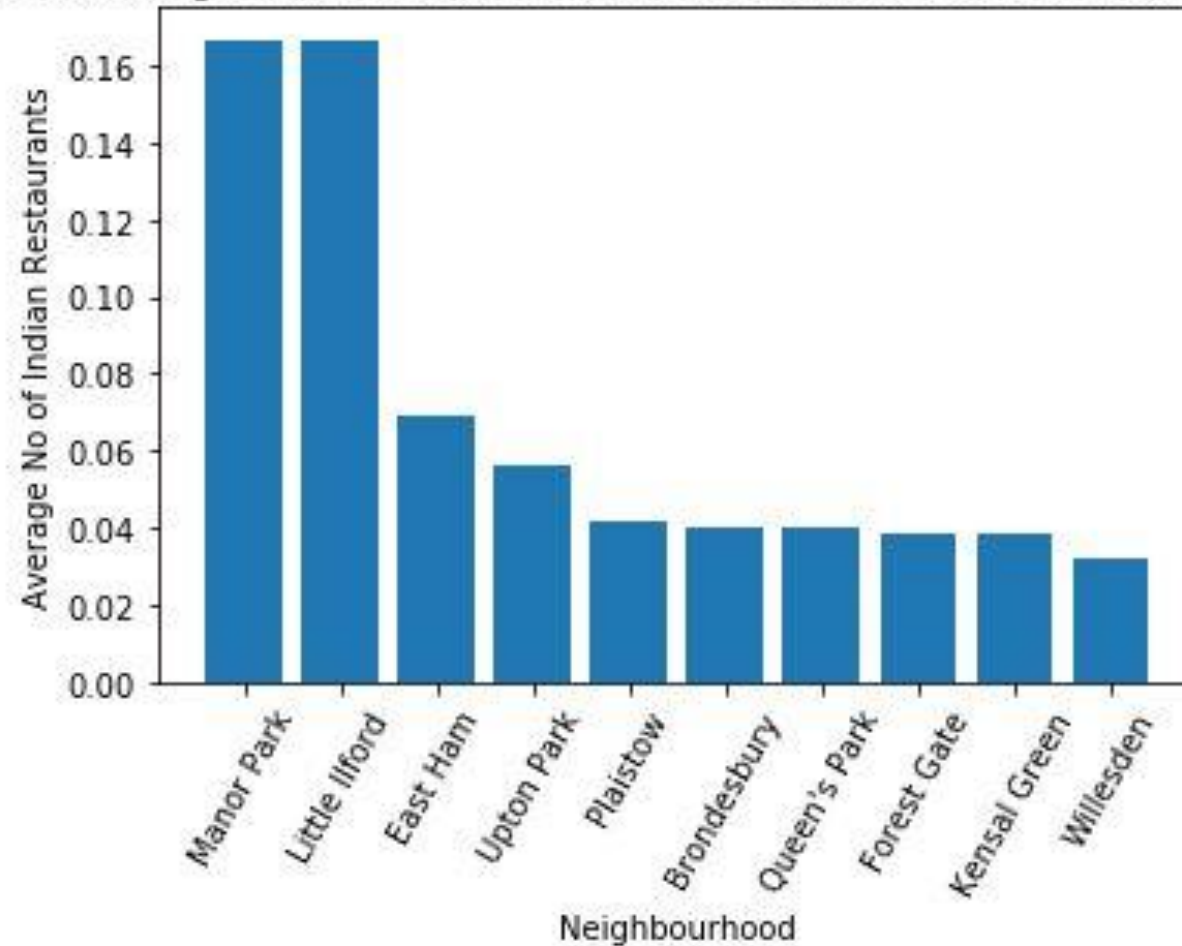
EXPLORATORY DATA ANALYSIS

- The number of Indian Restaurants in our relevant neighborhoods was 77, spread across different neighborhoods
- Preliminary analysis was done on Indian restaurants in our relevant neighborhoods
- Manor Park and Little Ilford have the largest average number of Indian Restaurants
- Large number of Indian Restaurants means large number of potential customers of Indian ethnicity



EXPLORATORY DATA ANALYSIS

Top 10 Average Number of Indian Restaurants across different Neighborhoods



CLUSTERING

- The top 10 most common venues in each neighborhood were gathered
- All the neighborhoods were combined with the columns of most common venues arranged from 1st to 10th
- Coffee Shops, Pubs and Cafes were the most common venues overall



CLUSTERING

	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	Beckton	Hotel	Grocery Store	Coffee Shop	Pub	Sandwich Place	Supermarket	Fast Food Restaurant	Pizza Place	Chinese Restaurant	Park
1	Bethnal Green	Coffee Shop	Pub	Café	Restaurant	Cocktail Bar	Wine Bar	Italian Restaurant	Yoga Studio	Chocolate Shop	Pilates Studio
2	Blackwall	Coffee Shop	Café	Italian Restaurant	Hotel	Park	Burger Joint	Sandwich Place	Plaza	Gym / Fitness Center	Bakery
3	Bow	Pub	Coffee Shop	Café	Grocery Store	Pizza Place	Bar	Convenience Store	Fast Food Restaurant	Art Gallery	Bus Stop
4	Brent Park	Platform	Grocery Store	Park	Pub	Plaza	Bus Stop	Brazilian Restaurant	Convenience Store	Middle Eastern Restaurant	Bus Line



KMEANS CLUSTERING

- For our analysis, we chose Kmeans Clustering since it is the most widely used and recognized clustering algorithm
- To find k (number of clusters), KElbowVisualizer package was used, which eases up the task of finding the best k for the data
- The number of clusters k was set to 4 (using elbow method)
- The neighborhoods were combined with their cluster labels and displayed in a dataframe

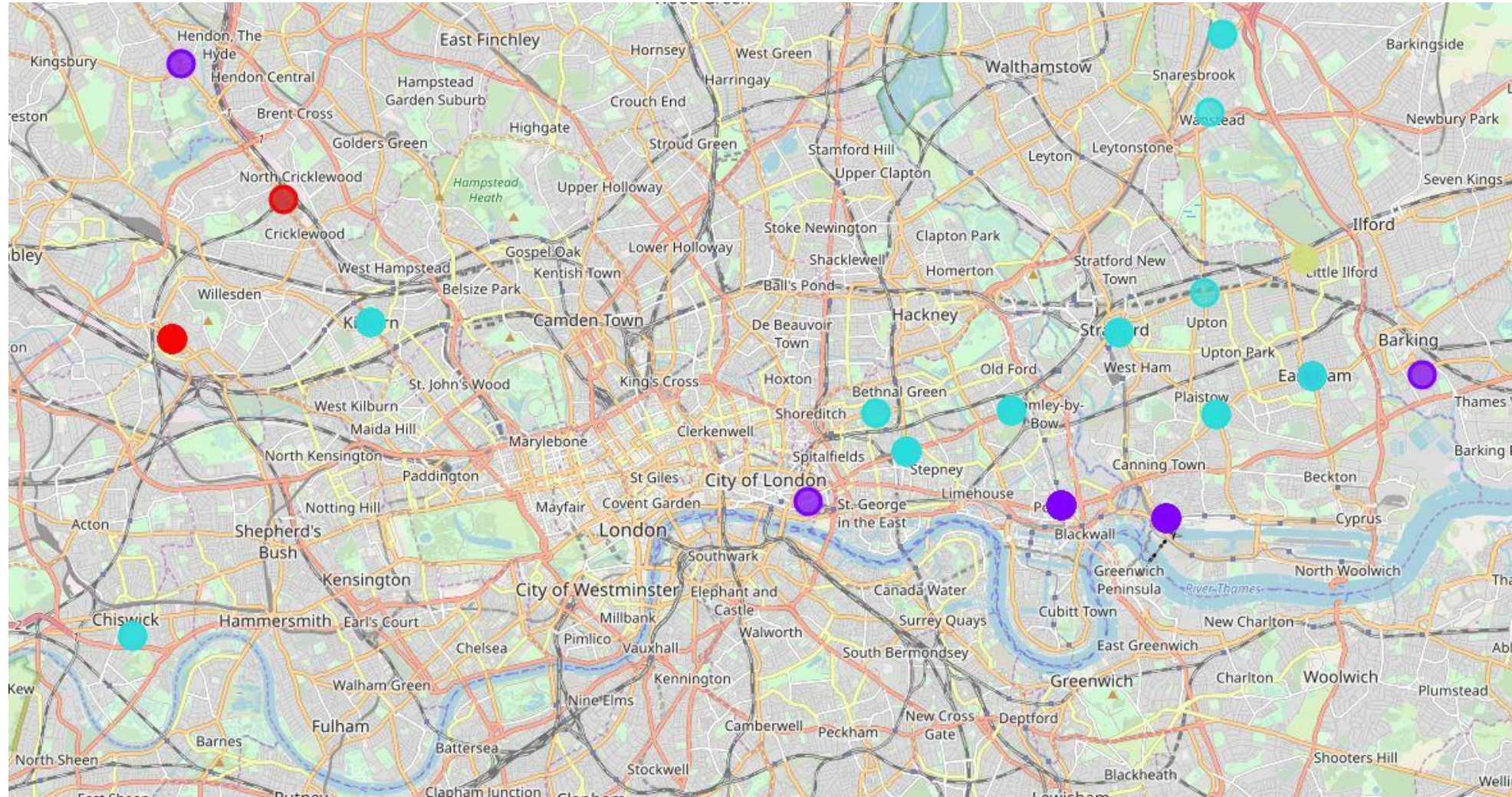


KMEANS CLUSTERING

	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	8th Most Common Venue	9th Most Common Venue
0	Beckton	Newham	E6	51.53292	0.054610	1	Hotel	Grocery Store	Coffee Shop	Pub	Sandwich Place	Supermarket	Fast Food Restaurant	Pizza Place	Chinese Restaurant
1	Beckton	Newham	E16	51.50913	0.015280	1	Hotel	Grocery Store	Coffee Shop	Pub	Sandwich Place	Supermarket	Fast Food Restaurant	Pizza Place	Chinese Restaurant
2	Beckton	Newham	IG11	51.53312	0.084077	1	Hotel	Grocery Store	Coffee Shop	Pub	Sandwich Place	Supermarket	Fast Food Restaurant	Pizza Place	Chinese Restaurant
3	Bethnal Green	Tower Hamlets	E2	51.52669	-0.062570	2	Coffee Shop	Pub	Café	Restaurant	Cocktail Bar	Wine Bar	Italian Restaurant	Yoga Studio	Chocolate Shop
4	Blackwall	Tower Hamlets	E14	51.51122	-0.012640	1	Coffee Shop	Café	Italian Restaurant	Hotel	Park	Burger Joint	Sandwich Place	Plaza	Gym/Fitness Centre



KMEANS CLUSTERING



RESULTS

- Cluster 4 (Little Ilford and Manor Park) was the ideal location for our business problem because of the number of Indian restaurants and hence the large customer base
- Cluster 3 (containing Queen's Park, Upton Park, Kensal Green, Eastham and Brondesbury) could also be considered because of the presence of good number of Indian Restaurants
- Cluster 1 and 2 had insignificant number of Indian Restaurants



DISCUSSION

- Little Ilford and Manor Park (in Newham borough) already have a lot of competition
- Queen's Park, Upton Park, Kensal Green and Brondesbury can be potential locations but all of them are in the Brent borough which has less proportion of Asian population than Newham, and hence lesser number of potential customers
- Eastham (located in Newham) has lesser number of Indian restaurants (less competition) but is in close proximity of Little Ilford and Manor Park, so it'll probably have almost the same number of potential customers
- Eastham also has a lot of shopping stores, grocery shops and bus stations so it can reach bigger number of potential customers



CONCLUSION

- Eastham is recommended as the target location
- Many other factors will have to be considered as well to make a final decision, such as:
 - Property rents
 - Ease of access
 - Traffic situation
 - Market structure



LIMITATIONS

- Foursquare's free developer account constraints
- Unavailability of price/cost related data of the neighborhoods
- Unavailability of the demographic data based on ethnicity
- Analyzing ethnic groups across different neighborhoods would result in much better insights



THANK YOU!

