



# THE BATTLE OF THE NEIGHBORHOODS

Sana Syed

IBM Data Science Capstone  
Project

# IDENTIFYING THE PRIME LOCATION FOR A THAI RESTAURANT IN ENGLAND

Chai San wants to open his first ever Thai Restaurant in England, however, having never visited the country before, he has no idea as to where would be the best location to begin looking in. Can we identify the best location for his new business venture?

# PROJECT OVERVIEW

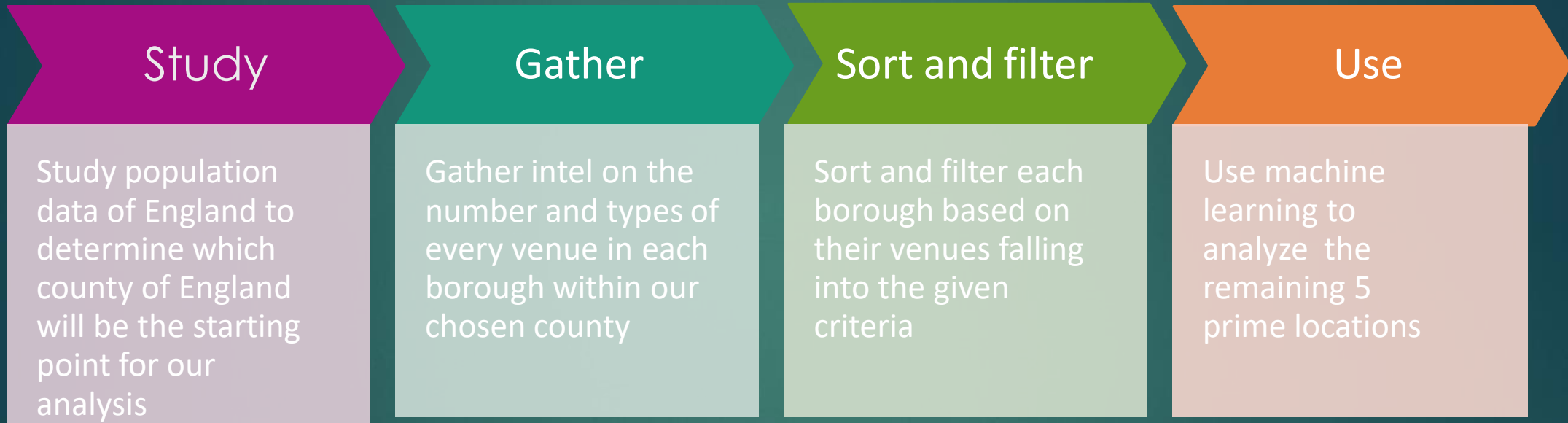
Identify a location for a new Thai restaurant in England

The location must be in an area with a high Asian population

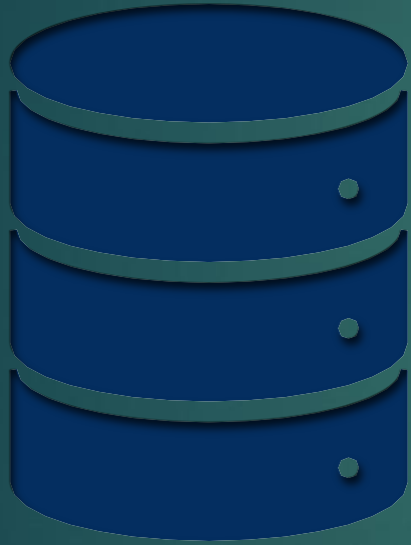
The location must not be in an area with other Thai restaurants in close proximity

The location must be in an area which is not densely populated with restaurants of any type

# METHODOLOGY



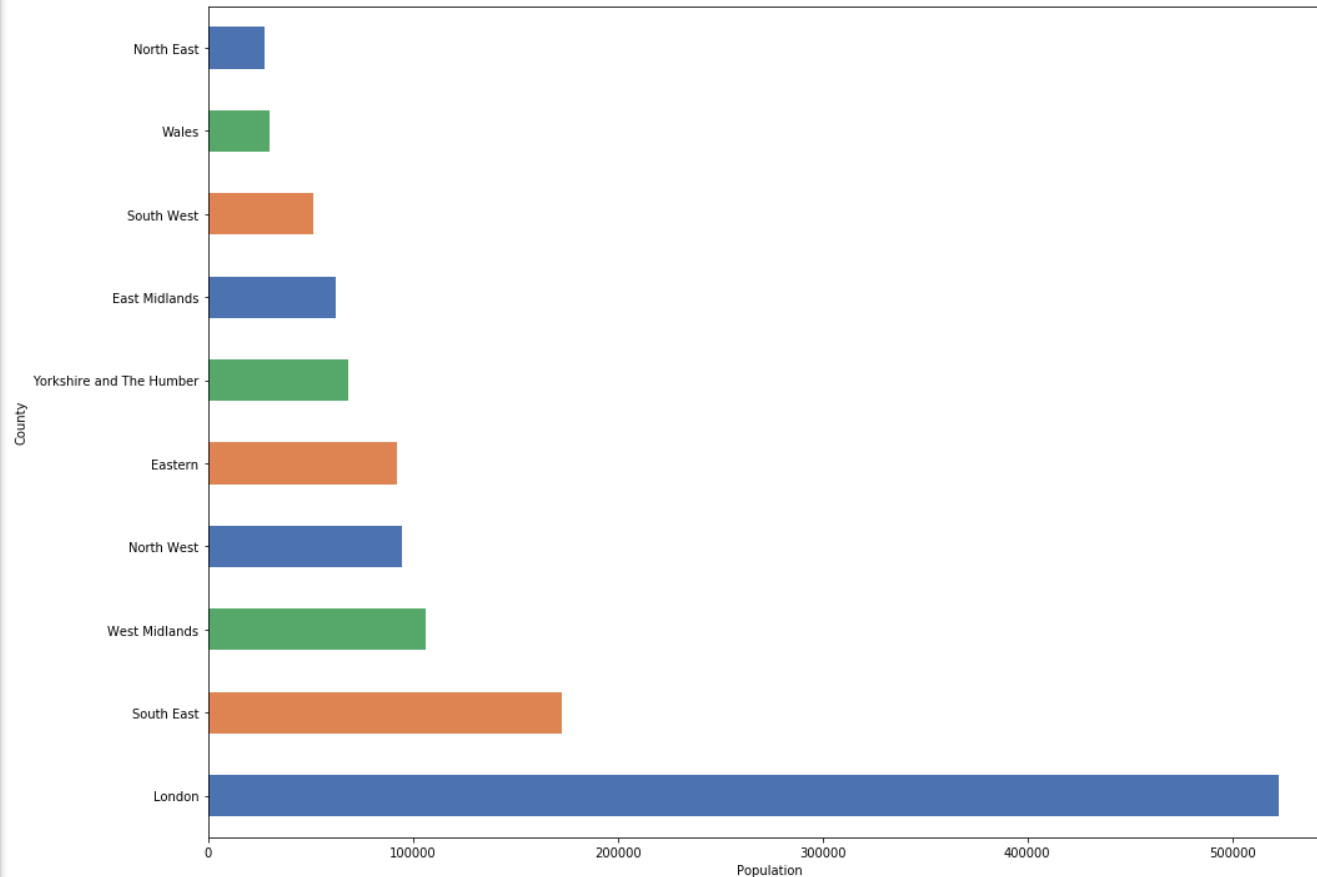
# DATA SOURCES



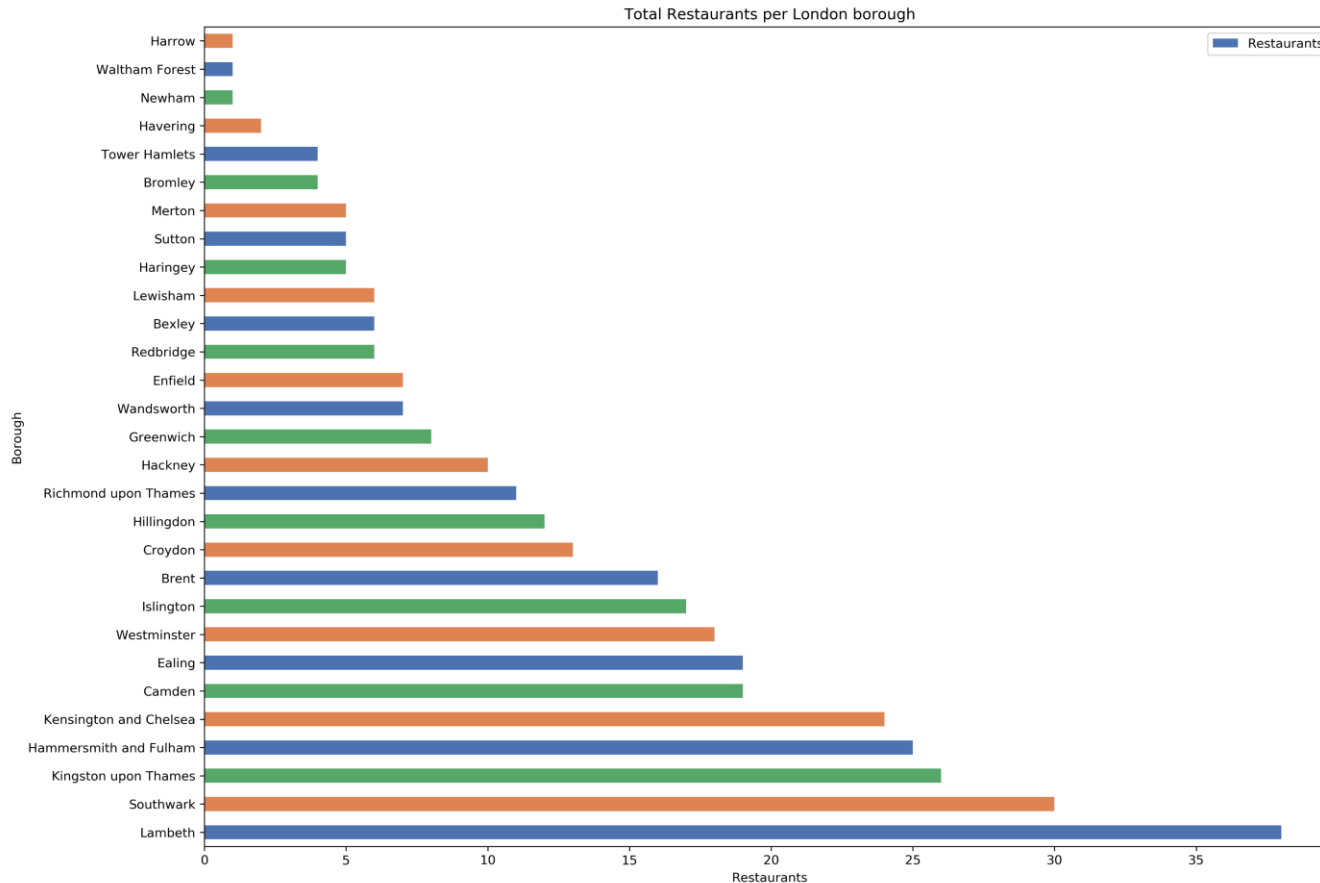
- Foursquare API for restaurant/venue information
- Office of National Statistics for Ethnic Populating

- ▶ Data
- ▶ □ <http://martinjc.github.io> For GeoJSON Data of England
- ▶ □ **Wikipedia** for a list of London boroughs and
- ▶ lat/lon values

# ASIAN POPULATION IN ENGLAND



□ Data shows that London has a far bigger Asian (Chinese or Other) population than anywhere else in the country with approximately 523,000 citizens. This meant that London was our starting point for further exploration.



## VENUE DISCOVERY

After querying each of the 32 boroughs of London using the Foursquare API, we identified and gathered a list of locations for all 1,477 venues in London and, more specifically, the locations of all Thai and non-Thai restaurants in London too. We found that there was a total of 348 restaurants in London, 16 of which are Thai restaurants.

# VENUE DISCOVERY (CONTINUED)

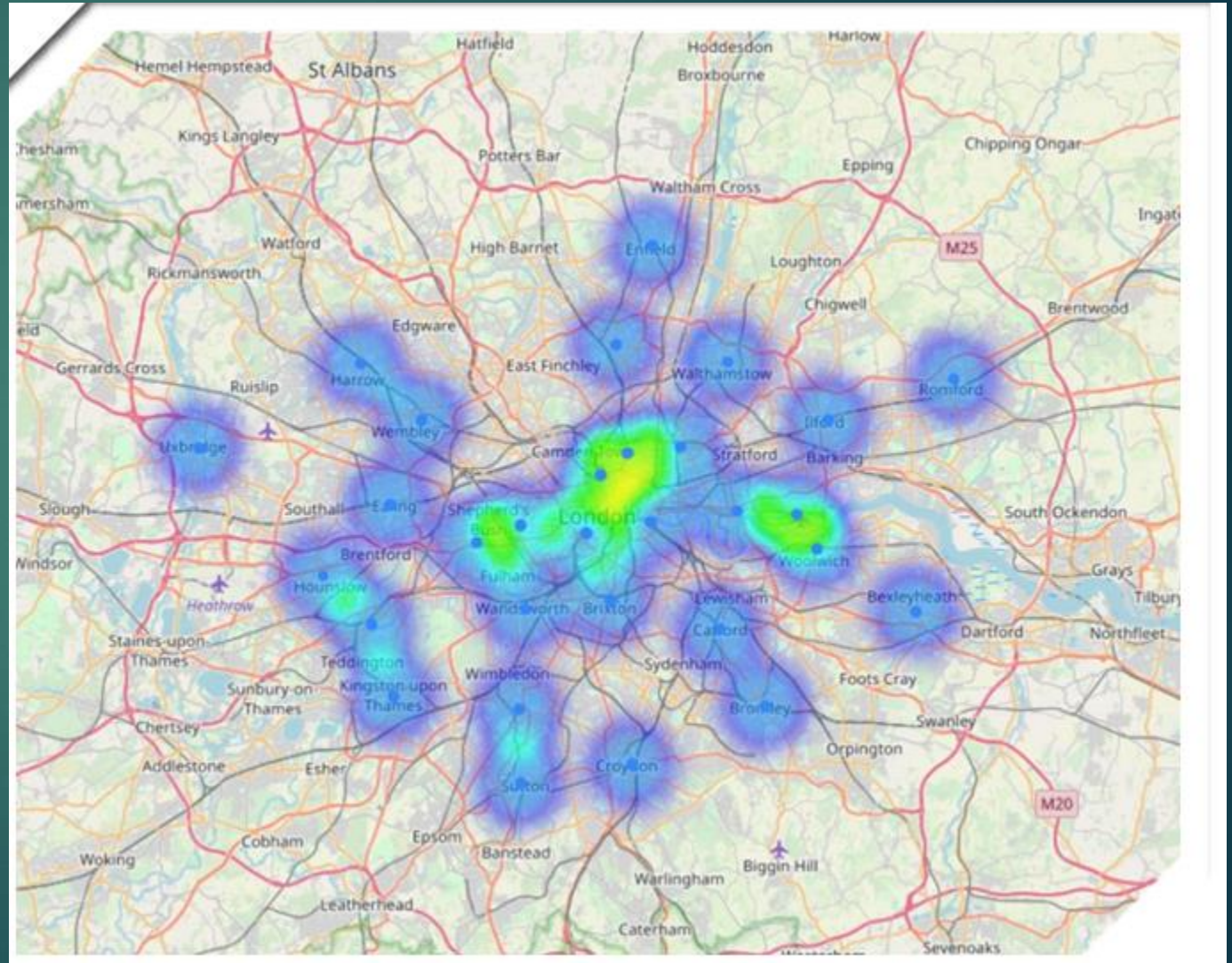
Barking and Dagenham has a total of 7 venues, including a Park, a Gym and a Supermarket.

	▼ Venue Name	▼ Cetegory	▼ Lat	▼ Long	▼
0	Central Park	Park	51.55956	0.161981	
1	Crowlands Heath Golf Course	Golf Course	51.562457	0.155818	
2	Robert Clack Leisure Centre	Martial Arts Dojo	51.560808	0.152704	
3	Beacontree Heath Leisure Centre	Gym / Fitness Center	51.560997	0.148932	
4	Morrisons Becontree Heath	Supermarket	51.559774	0.148752	
5	Becontree Heath Bus Station	Bus Station	51.561065	0.150998	
6	Dagenham Swimming Pool	Pool	51.560946	0.150054	



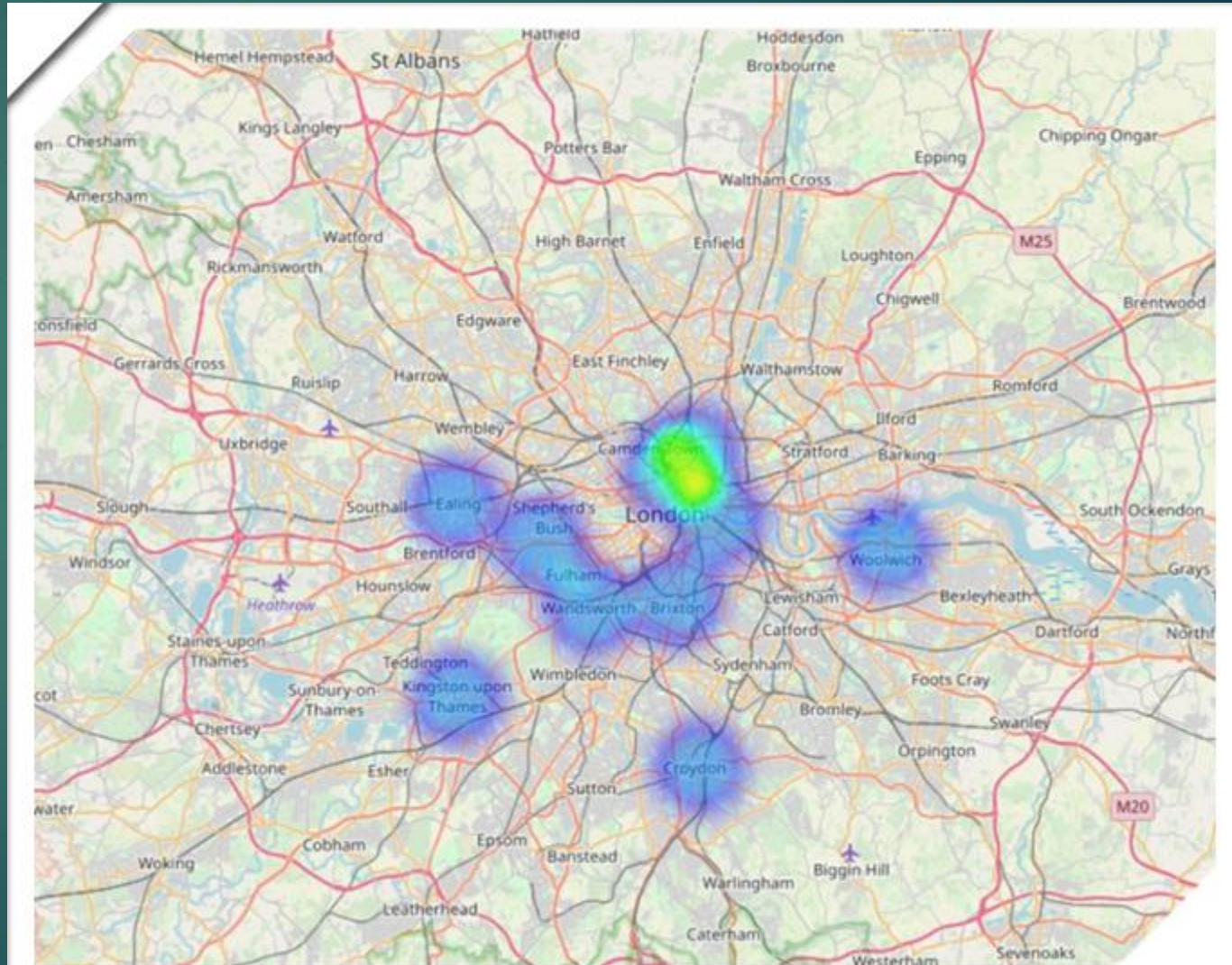
# RESTAURANTS IN LONDON

□ Using the data we could generate a heat-map of all restaurants in London to determine where the “hotspots” are.



# THAI RESTAURANTS IN LONDON

□ Using the same data we could also generate a heat-map of all Thai restaurants in London to determine where the “hotspots” are.





# MACHINE LEARNING

□ Once the dataset was adequately filtered, machine learning was used to cluster each of the remaining 5 boroughs based on their similarity

Borough	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
Waltham Forest	0	Grocery Store	Pub	Gym	Beer Store	Gym / Fitness Center
Tower Hamlets	1	Hotel	Italian Restaurant	Coffee Shop	Pizza Place	Sandwich Place
Bromley	2	Coffee Shop	Clothing Store	Burger Joint	Pizza Place	Gym / Fitness Center
Havering	2	Clothing Store	Coffee Shop	Shopping Mall	Café	Fast Food Restaurant
Newham	3	Hotel	Airport	Duty-free Shop	Chinese Restaurant	Light Rail Station

# RESULTS

Our study managed to accurately find and select a series of locations which meet the business problem given by our client.

5 boroughs which had a high Asian population, the lowest restaurant density and also contained zero Thai restaurants were identified and probed.

# CONCLUSION

- Population data of ethnic groups within the UK accurately identified London as the county with the highest Asian population, therefore London was selected for further probing.
- Within London, each borough was analysed based on the number of venues, number of restaurants and number of Thai restaurants that they each contained.
- These boroughs were then sorted and filtered based on the criteria given to leave just 5 boroughs which could be labelled as prime locations for the new restaurant.
- Machine Learning was applied to the 5 selected boroughs to cluster each borough based on similarity and their 5 most common venues.
- These 5 boroughs could now be put-forward to the client for his consideration and, if the client so wishes, each borough could be probed further using the same techniques on a more concentrated level.