Starting a Restaurant in London, UK

a Data Science Project

Introduction

- We have been approached by a company interested in investing in a new restaurant opening within London.
- They have asked us to determine where would be the best location and what sort of restaurant would be most appropriate for that area.
- The business problem and purpose of this project therefore is to select a borough in London to open a new restaurant and select its cuisine. In order to optimize the location and restaurant type we will characterize the boroughs of London using k-means clustering.
- The people interested in our results will be those looking to invest in restaurant within London.

Data

- For the borough data we are scraping the webpage: https://en.wikipedia.org/wiki/List_of_London_boroughs
- This data will be supplemented with data on up to the top 100 venues within a 1km area the London borough. This additional data will be acquired from Foursquare.

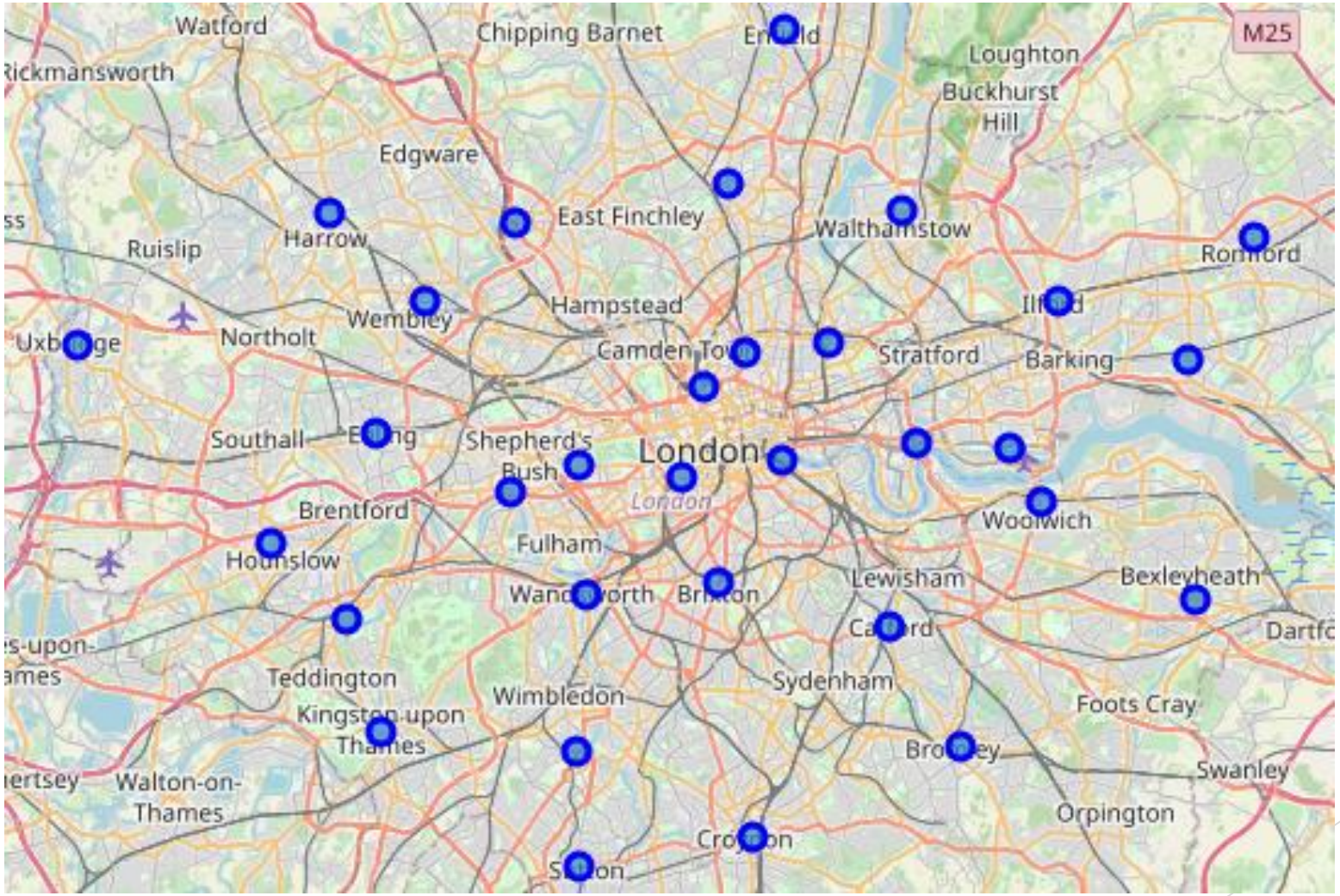
Borough Data

- To begin with we process the scraped wiki data to get a dataframe which looks like the following;

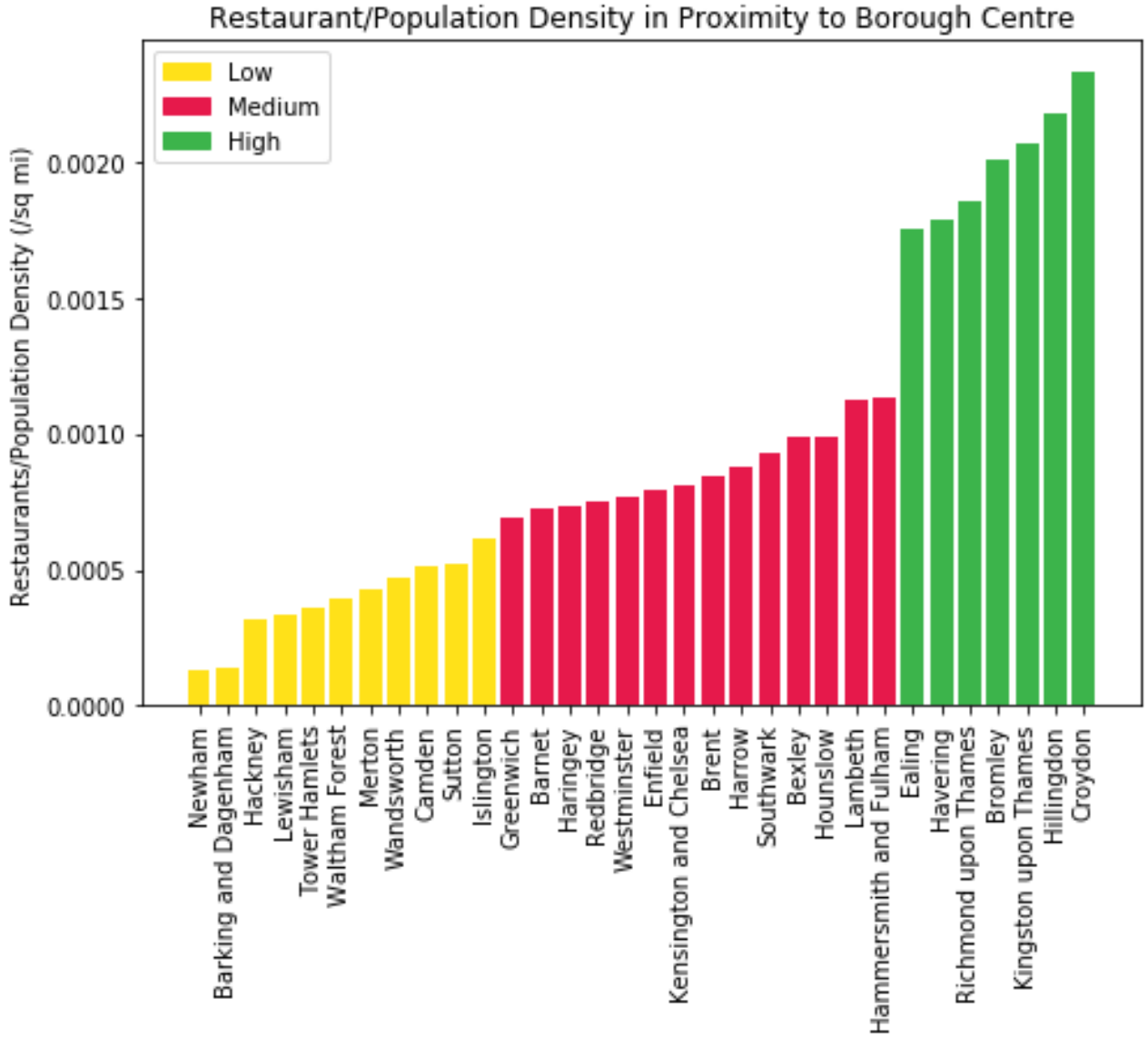
	Borough	Population Density (sq mi)	Latitude	Longitude
0	Barking and Dagenham	13952.0	51.5607	0.1557
1	Barnet	11021.0	51.6252	-0.1517

- The dataframe contains information on all the London boroughs.

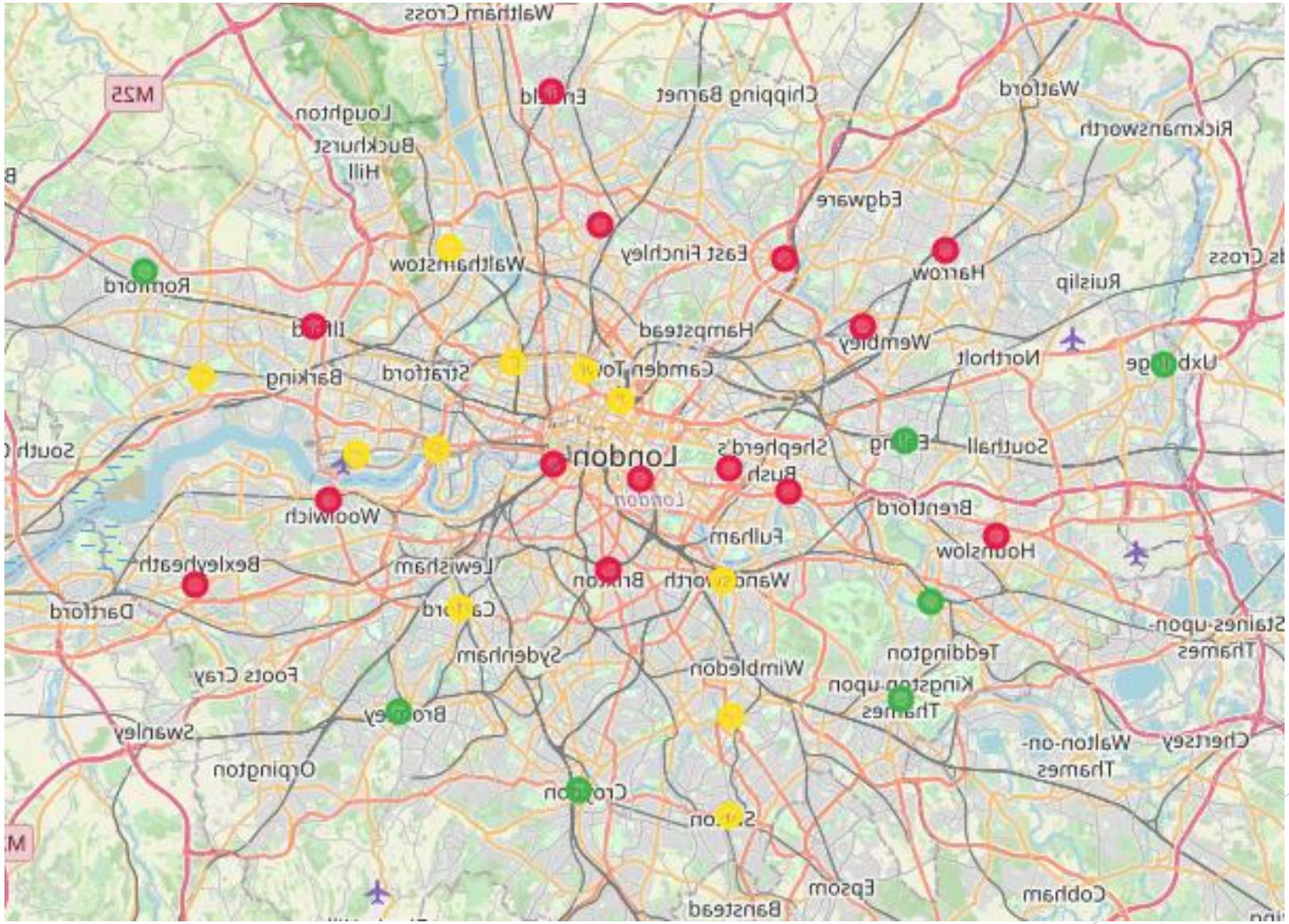
Borough Data



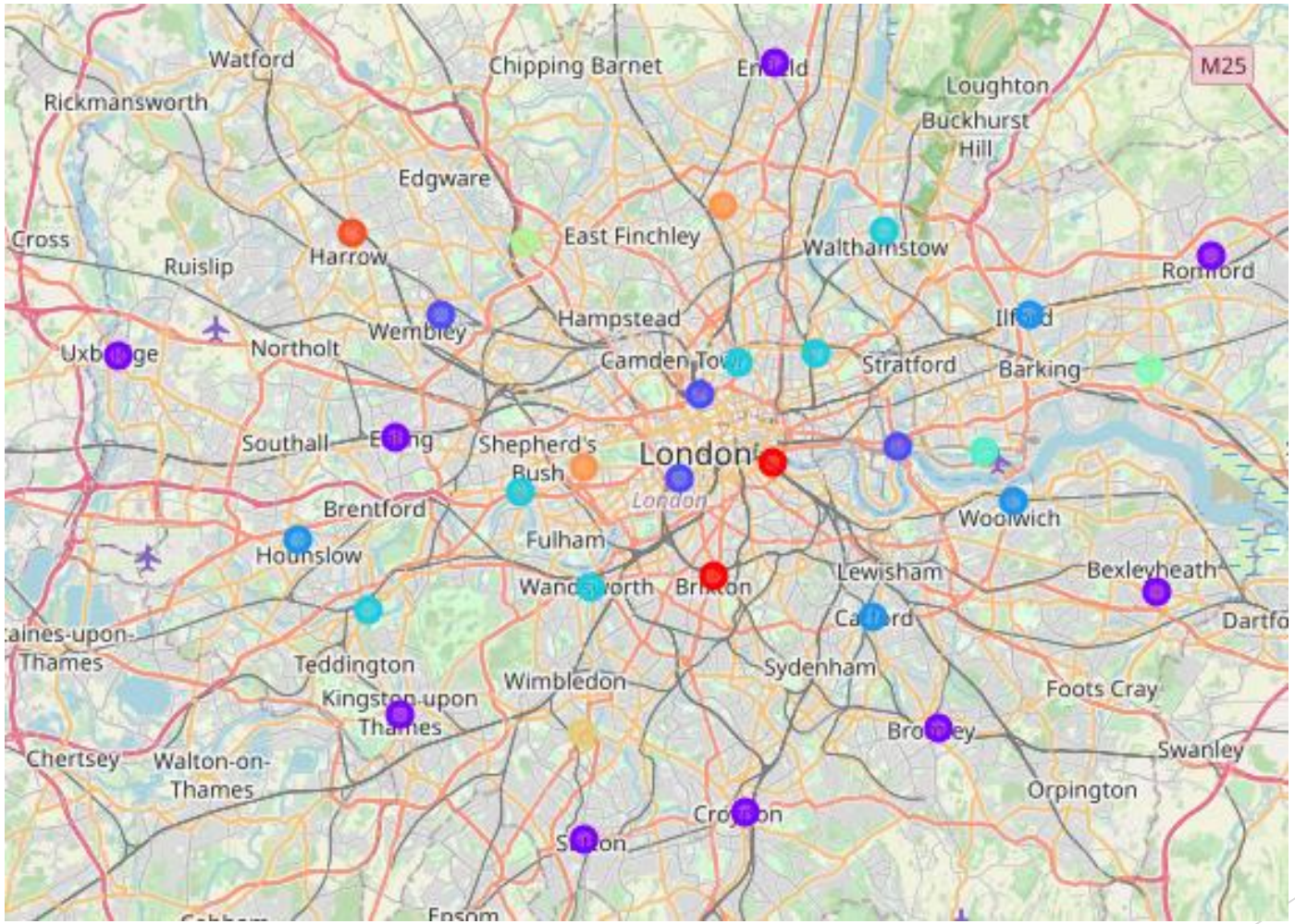
Clustering based on Restaurant/Population Density



Clustering based on Restaurant/Population Density



Clustering based on Borough Features (ex. restaurants)



Feature Cluster 1

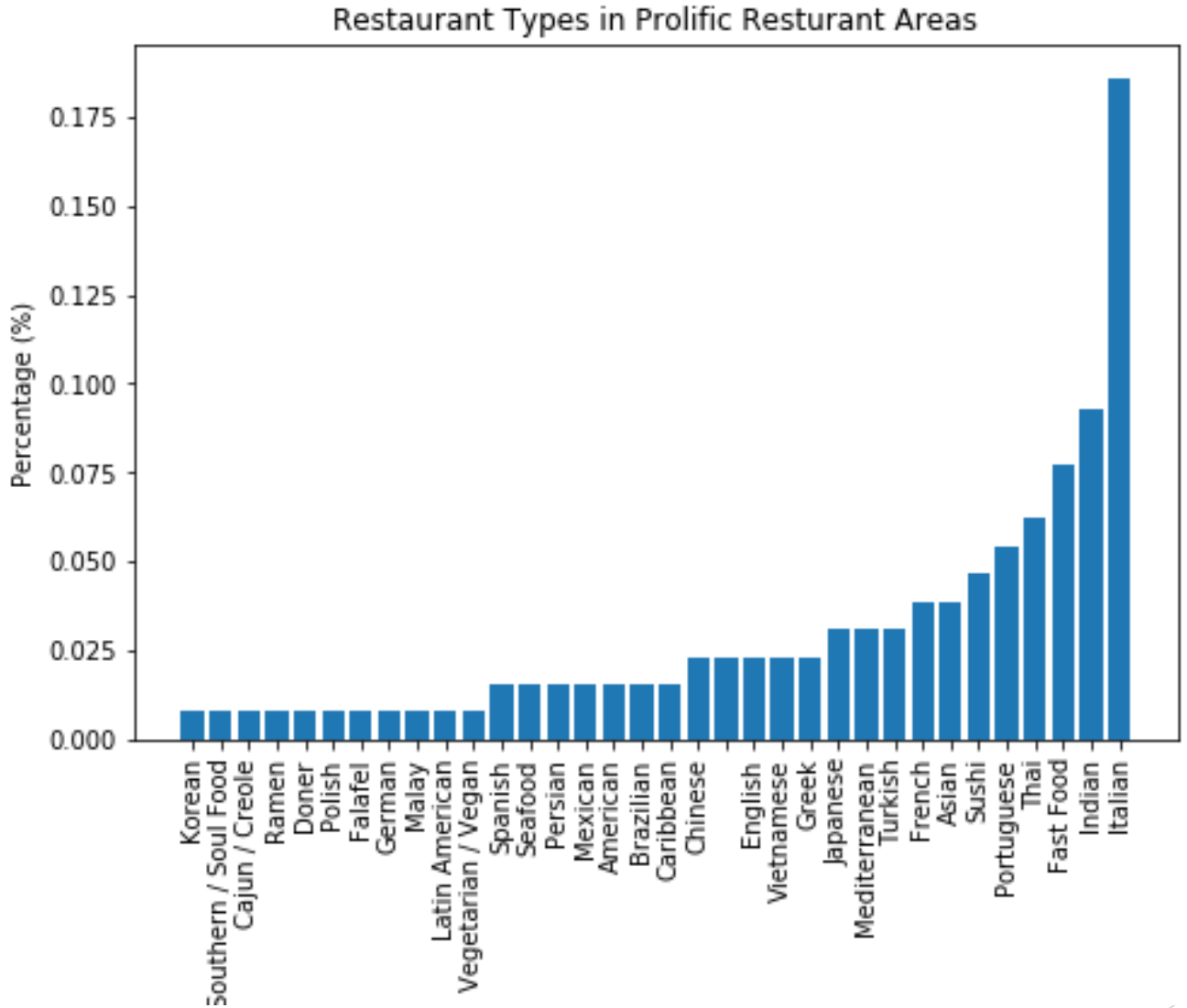
Borough	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
Bexley	Pub	Clothing Store	Coffee Shop	Supermarket	Hotel	Pharmacy	Furniture / Home Store	Nightclub	Discount Store	Department Store
Bromley	Pub	Clothing Store	Coffee Shop	Gym / Fitness Center	Supermarket	Electronics Store	Pizza Place	Burger Joint	Café	Stationery Store
Croydon	Coffee Shop	Pub	Hotel	Clothing Store	Bookstore	Café	Sandwich Place	Platform	Gym / Fitness Center	Park
Ealing	Coffee Shop	Pub	Pizza Place	Bakery	Park	Burger Joint	Hotel	Café	Gym / Fitness Center	Supermarket
Enfield	Pub	Clothing Store	Coffee Shop	Grocery Store	Supermarket	Optical Shop	Shopping Mall	Gift Shop	Pharmacy	Café
Havering	Coffee Shop	Pub	Clothing Store	Shopping Mall	Park	Furniture / Home Store	Café	Supermarket	Grocery Store	Department Store
Hillingdon	Coffee Shop	Pub	Clothing Store	Pharmacy	Gym	Supermarket	Bar	Burger Joint	Bookstore	Park
Kingston upon Thames	Coffee Shop	Pub	Café	Burger Joint	Park	Clothing Store	Department Store	Hotel	Bookstore	Bar
Sutton	Coffee Shop	Clothing Store	Pub	Café	Hotel	Bar	Sandwich Place	Department Store	Pizza Place	Supermarket

- This cluster has ~85% of the boroughs in the relatively high number of restaurant clusters, indicating that these shared features correlate with restaurant success.

Selecting the best Borough

- Boroughs in the high restaurant feature cluster: Bexley, Bromley, Croydon, Ealing, Enfield, Havering, Hillingdon, Kingston upon Thames, Sutton
- Boroughs with a low restaurant industry relative to the population density are: Newham, Barking and Dagenham, Hackney, Lewisham, Tower Hamlets, Waltham Forest, Merton, Wandsworth, Camden, Sutton, Islington
- We can then look for overlap between the two clusters and select the borough with the relatively lowest number of restaurants:
- The boroughs with features consistent with a successful restaurant industry but a low restaurant presence are: **Sutton**

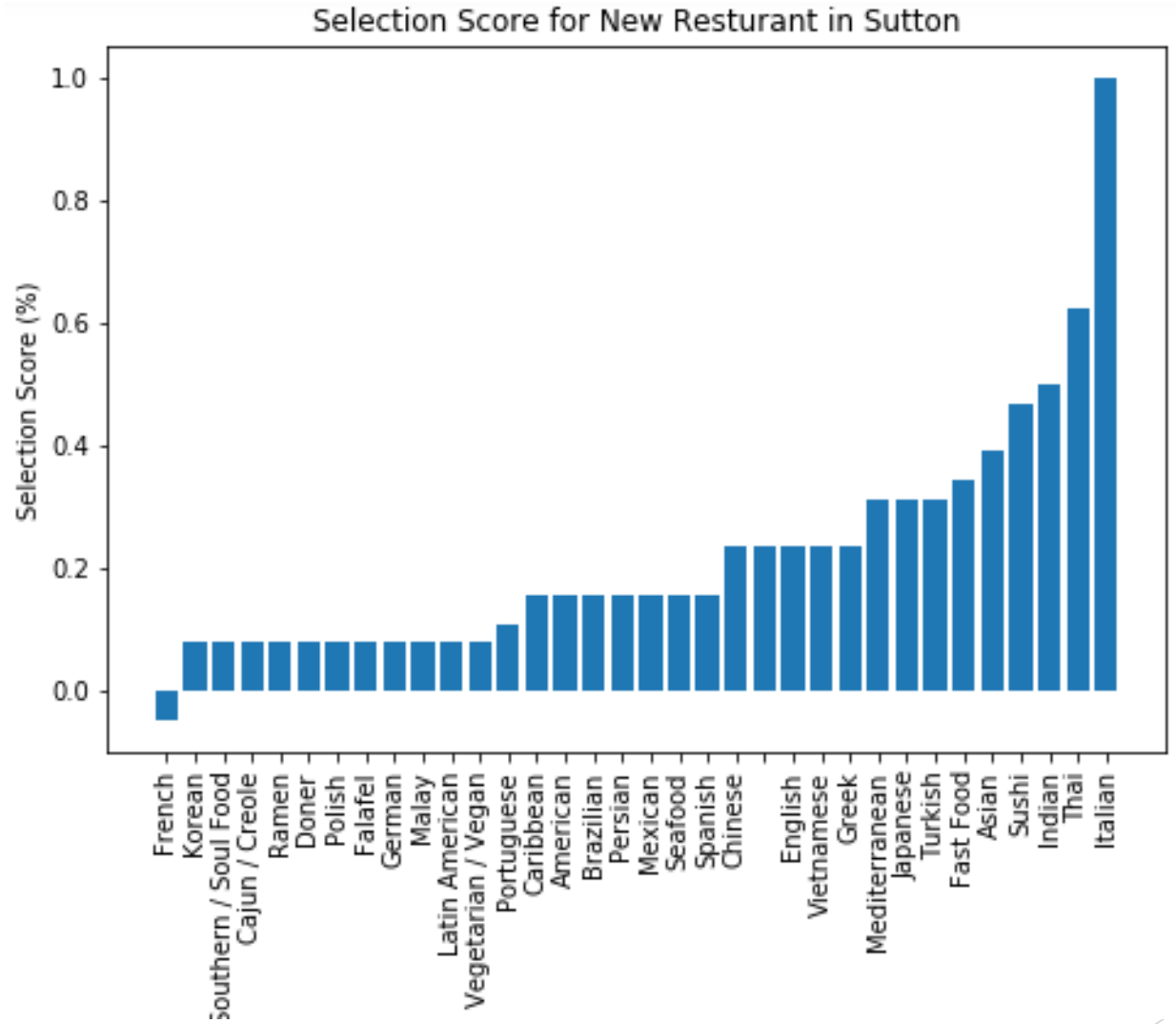
Selecting the Best Restaurant Cuisine



Selecting the Best Restaurant Cuisine

	Expected number	True number	Difference
Sushi Restaurant	1.0	0.0	1.0
Portuguese Restaurant	1.0	1.0	0.0
Thai Restaurant	1.0	0.0	1.0
Fast Food Restaurant	1.0	1.0	0.0
Indian Restaurant	2.0	1.0	1.0
Italian Restaurant	4.0	2.0	2.0

Selecting the Best Restaurant Cuisine



Discussion

- Based on the analysis above we identified the most promising borough as Sutton. Sutton shares feature with many of the boroughs which have high population of restaurants relative to the population density, but Sutton has very few restaurants itself given its population density.
- Sutton was the only low restaurants borough in the feature cluster containing ~85% of the high restaurant boroughs.
- Having identified Sutton we developed a scoring system to determine the type of restaurants that would be the most successful. Based on this analysis the top two cuisines were Italian followed by Thai. It should be noted however that there are already two Italian restaurants in Sutton, but no Thai.
- We therefore recommend Thai as based on the available data from similar boroughs will be a high demand for Thai cuisine and there is at present no direct competitor.

Conclusions

- Borough: Sutton
- Restaurant: Thai
- We might also want to consider whether:
  - Areas might have a low population but a high degree of footfall.
  - The location of boroughs may not be at major towns.
  - Whether a finer granularity at the scale of neighbourhoods would give more reliable results.