THE BATTLE OF NEIGHBOURHOODS

Finding an ideal location of opening a healthy food and supplements store in London

Introduction

- Object: to define some ideal locations for opening a healthy food and supplements store in London, based on venues data of each neighborhood , for a well-known Asian firm
- Target location: obviously, near gyms and fitness centers, as this portion of people visiting these places is more prone to buy company's providing products

Data

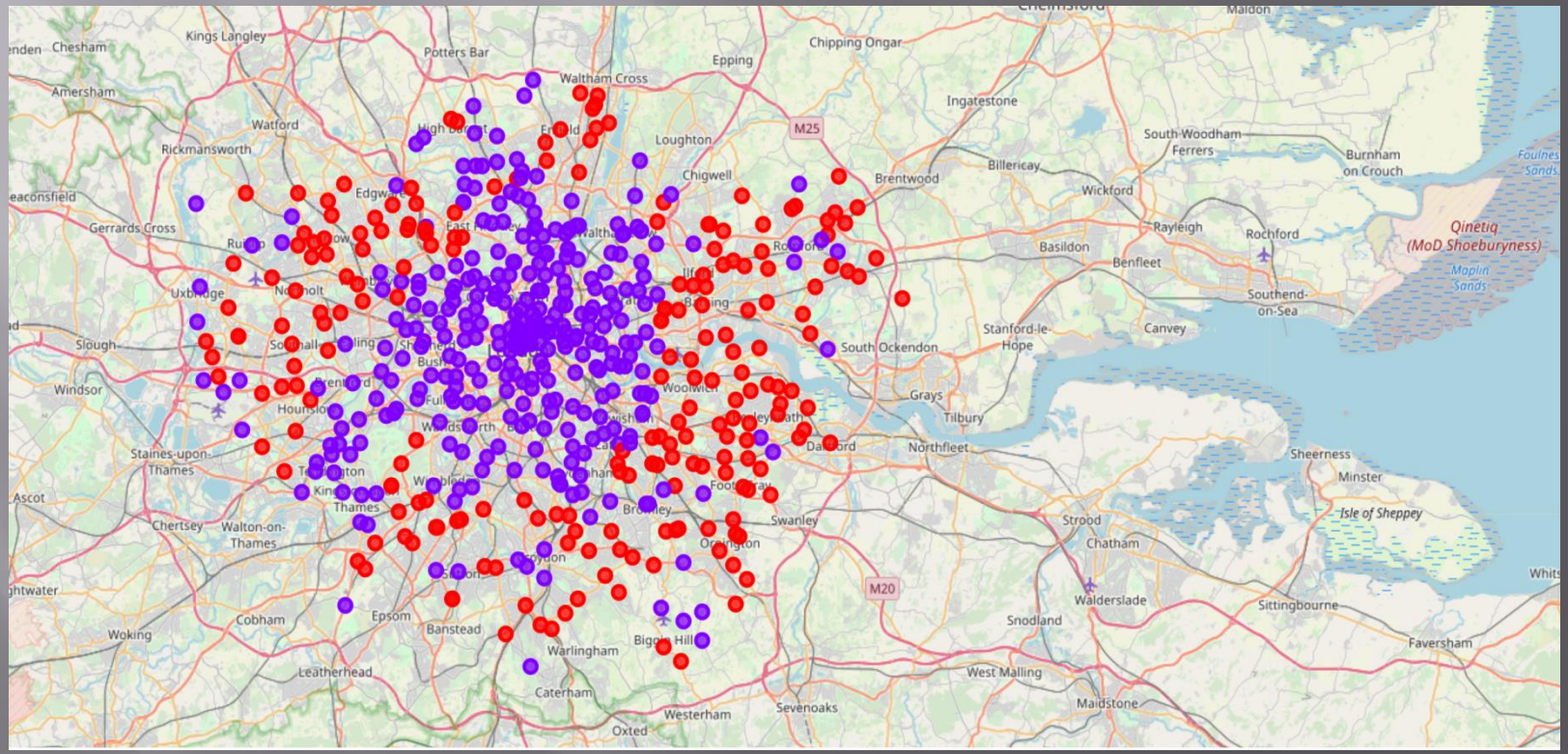
- Data from wikipedia to get London's neighborhoods
- Data from Google Maps geocoding API to get the exact location(latitude and longitude) of each neighborhood
- Data from Foursquare to get nearest venues of each neighborhood

Methodology

- Collecting all appropriate data and pre-processing
- Find 10 most common venues for each neighborhood
- Cluster neighborhoods using K-Means algorithm , having found first optimal K
- Examine the algorithm's results and suggest solutions

Results(1)

- 2 clusters, containing similar venues



Results(2)

- 6 different ideal locations, 3 on each cluster, having as 1st common venue Gym / Fitness Center

	Location	Dial code	OS grid ref	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	10th Most Common Venue
73	Burroughs, The	020	TQ227891	51.587840	-0.229035	1	Gym / Fitness Center	Pub	Coffee Shop	Park	Café	Department Store	History Museum	Grocery Store	Pizza Place	Golf Course
77	Canary Wharf	020	TQ375802	51.505431	-0.023533	1	Gym / Fitness Center	Coffee Shop	Plaza	Park	Bakery	Burger Joint	Hotel	Ramen Restaurant	Pizza Place	Bar
105	Colindale	020	TQ213897	51.589691	-0.249103	0	Gym / Fitness Center	Supermarket	Coffee Shop	Pub	Hotel	Park	Fast Food Restaurant	Grocery Store	History Museum	Chinese Restaurant
253	The Hyde	020	TQ215888	51.586400	-0.249812	0	Gym / Fitness Center	Pub	Park	Supermarket	Coffee Shop	Pet Store	Hotel	History Museum	Chinese Restaurant	Korean Restaurant
351	Old Oak Common	020	TQ216823	51.520470	-0.250717	1	Gym / Fitness Center	Middle Eastern Restaurant	Pub	Park	Gastropub	Thai Restaurant	Coffee Shop	Office	Fast Food Restaurant	Grocery Store
479	Upper Ruxley	020, 01689	TQ4970	51.412110	0.151023	0	Gym / Fitness Center	Golf Course	Italian Restaurant	Restaurant	Convenience Store	Event Service	Coffee Shop	Fast Food Restaurant	Park	Paintball Field

Discussions

- 2 clusters, based on venues in London => all venues are located similarly on each neighborhood => London a modern city
- By adding more data(-constraints) like neighborhood's crime rates(-safety), average income(quality of target group) and housing prices(-rental prices),etc we can build a more accurate model
- Generally speaking, the more **DATA** we have **is POWER** the more accurate our model IS

Conclusion

- Goal: find some ideal locations, based on neighborhood's venues occurancy
- Data: collect, combine and process data from wikipedia, Google Maps geocoding API and Foursquare
- Use of K-Means algorithm to segment neighborhoods by most common venues and find where target places are
- Remember: Data is POWER. By collecting, combing and analyzing data, we can build great models, achieving great results!

Thank you