The Battle of Neighborhoods

"Wontons in London"

IBM Capstone Project
2020

The Challenge

To find the most suitable location for a new Chinese restaurant in London, UK.

To best satisfy the following conditions:

- Affordable
- Competitive
- Busy

Can we use *data science* to complete this task?

Methodology

Resources:

- Wikipedia
- Property Data
- Foursquare Location Data

Techniques:

- Exploratory Analysis
- Machine Learning

Process:

- Acquire London Property Pricing Data
- Acquire and Explore London Population Data
- Suggest Locations

London Property Pricing Data

	Location	London borough	Post town	Postcode district	Dial Code	OS grid ref
0	Abbey Wood	Bexley	LONDON	SE2	020	TQ465785
1	Acton	Ealing	LONDON	W3	020	TQ205805
2	Addington	Croydon	CROYDON	CR0	020	TQ375645
3	Addiscombe	Croydon	CROYDON	CR0	020	TQ345665
4	Albany Park	Bexley	BEXLEY	DA5	020	TQ478728

+

	Area	Avg yield	Avg price	£/sqft	5yr +/-	Explore data
0	BR1	3.3%	£431,123	£461	+23%	Explore data
1	BR2	3.5%	£489,265	£467	+24%	Explore data
2	BR3	3.6%	£440,455	£489	+24%	Explore data
3	BR4	-	£628,666	£451	+21%	Explore data
4	BR5	2.9%	£450,523	£423	+23%	Explore data

=

	Location	London borough	Post town	Postcode district	Dial Code	OS grid ref	Avgprice
120	Blackfriars	City	LONDON	EC4	020	TQ318808	£1,016,190
121	Temple	City	LONDON	EC4	020	TQ311809	£1,016,190
134	Bloomsbury	Camden	LONDON	WC1	020	TQ305825	£1,031,839
135	Holborn	Camden	LONDON	WC1	020	TQ305815	£1,031,839
136	King's Cross	Camden and Islington	LONDON	WC1	020	TQ315835	£1,031,839
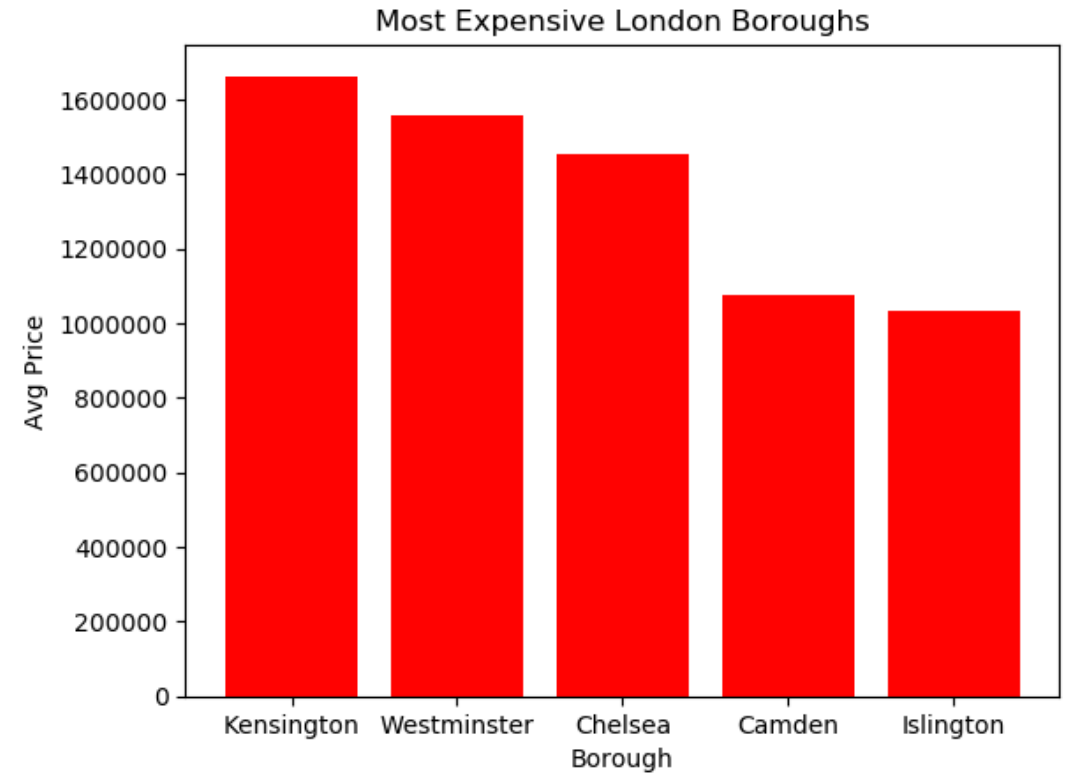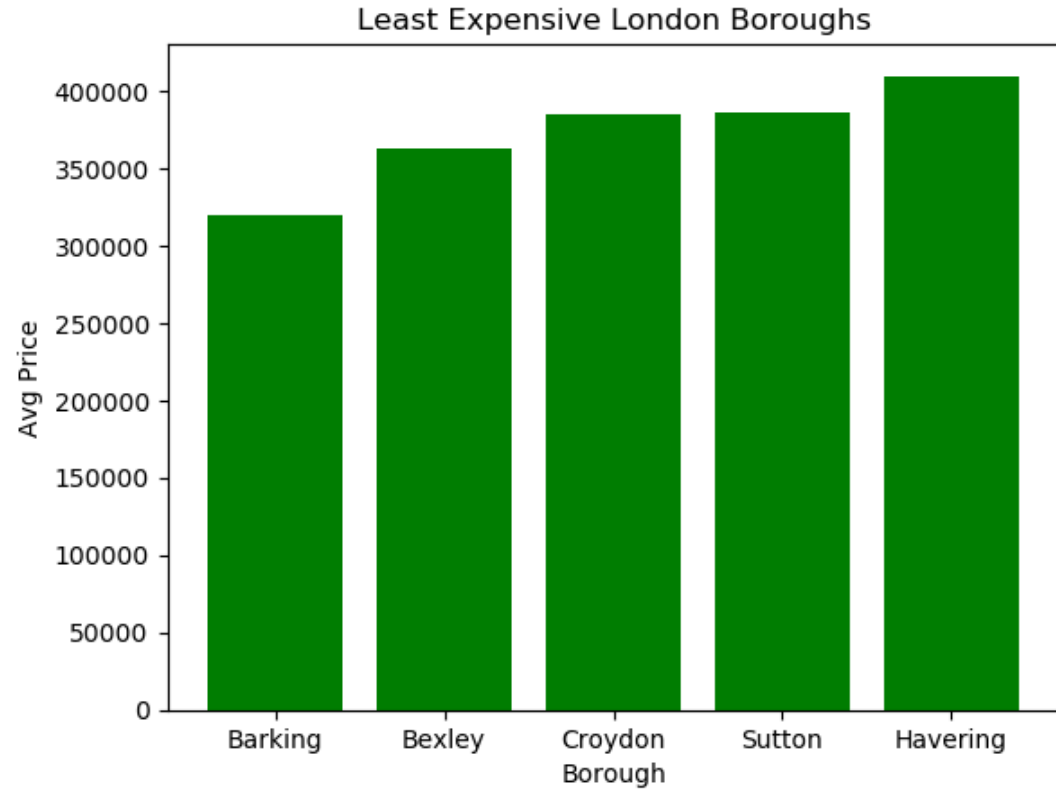
Python matplotlib function

London Population Data

	Rank	London Borough	Indian Population	Pakistani Population	Bangladeshi Population	Chinese Population	Other Asian Population	Total Asian Population
0	1	Newham	42484	30307	37262	3930	19912	133895
1	2	Redbridge	45660	31051	16011	3000	20781	116503
2	3	Brent	58017	14381	1749	3250	28589	105986
3	4	Tower Hamlets	6787	2442	81377	8109	5786	104501
4	5	Harrow	63051	7797	1378	2629	26953	101808
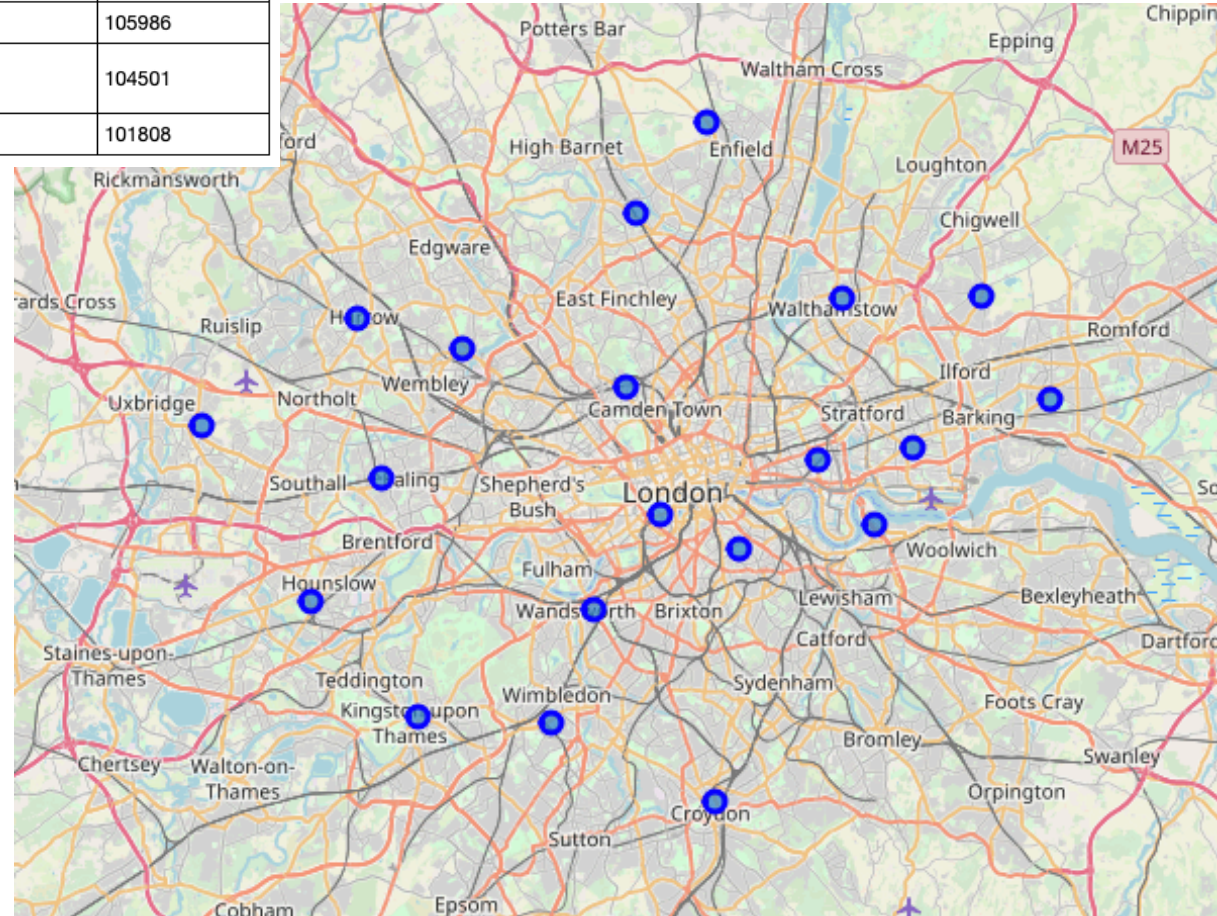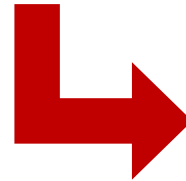
Python folium function

|	Borough	Borough Latitude	Borough 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	West Ham Food Centre	51.527451	0.031644	Convenience Store

One hot encoding
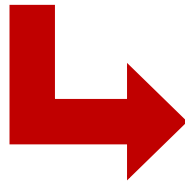
Foursquare API

	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
0	Barking and Dagenham	Construction & Landscaping	Lake	Park	Women's Store	Fish & Chips Shop	Department Store	Dessert Shop
1	Barnet	Bus Stop	Home Service	Café	Rental Car Location	Women's Store	Electronics Store	Fast Food Restaurant
2	Brent	Middle Eastern Restaurant	IT Services	Bus Station	Fast Food Restaurant	Food Truck	Supermarket	Café
3	Camden	Café	Gym / Fitness Center	Coffee Shop	Gastropub	Pizza Place	Bakery	Museum

|	Rank	London Borough	Indian Population	Pakistani Population	Bangladeshi Population	Chinese Population	Other Asian Population	Total Asian Population	Borough	1st Most Common Venue	2nd Most Common Venue
0	1	Newham	42484	30307	37262	3930	19912	133895	Newham	Café	Grocery Store
1	2	Redbridge	45660	31051	16011	3000	20781	116503	Redbridge	Indian Restaurant	Super...
2	3	Brent	58017	14381	1749	3250	28589	105986	Brent	Middle Eastern Restaurant	IT Serv...
3	4	Tower Hamlets	6787	2442	81377	8109	5786	104501	Tower Hamlets	Pizza Place	Canal L...
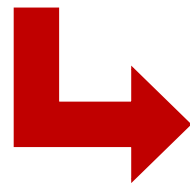
Clustering results
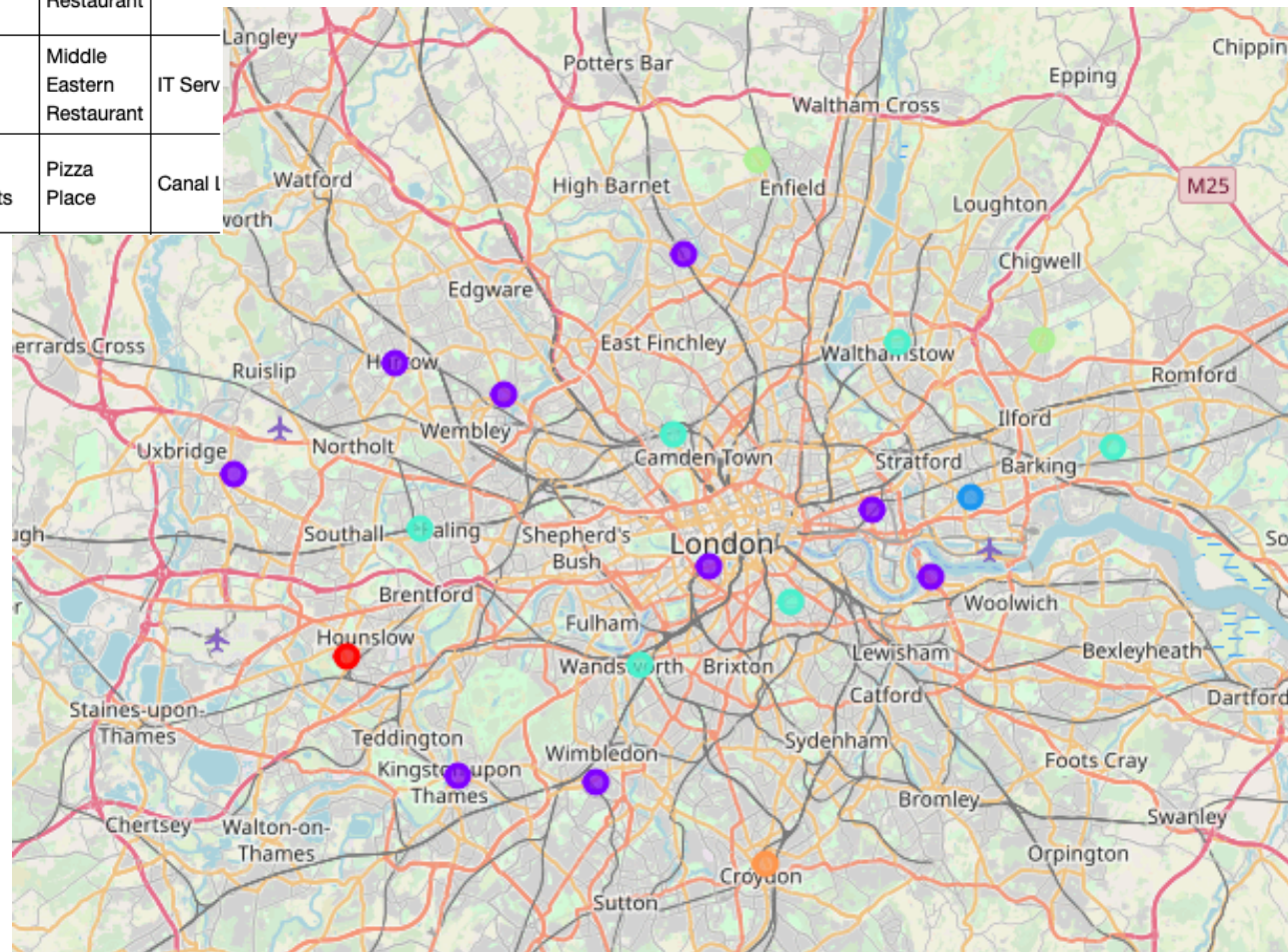
|	london_borough	avgprice	rank	indian_population	pakistani_population	bangladeshi_population	chinese_population
0	Barking and Dagenham	319211	18	7436	8007	7701	1315
1	Croydon	384420	10	24660	10865	2570	3925
2	Hounslow	474877	7	48161	13676	2189	2405
3	Enfield	485936	14	11648	2594	5599	2588
4	Ealing	504298	6	48240	14711	1786	4132
5	Redbridge	446953	2	45660	31051	16011	3000
6	Hillingdon	435438	8	36795	9200	2639	2889

Merged datasets
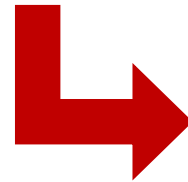
Scatter Plot
Average Property Price Vs Chinese Population
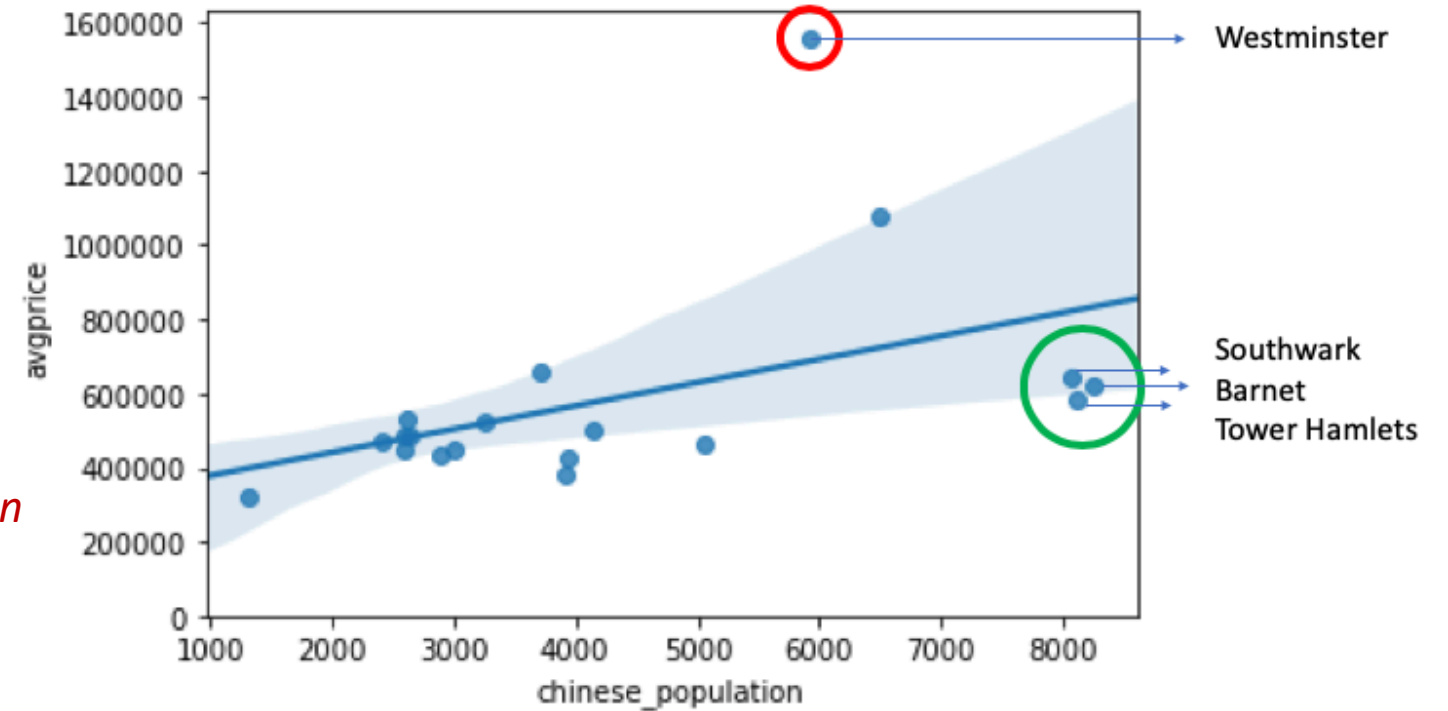
Suggested Locations

- Barnet
- Tower Hamlets
- Southwark

Foursquare "Venues"
Local competition
(5000m radius)
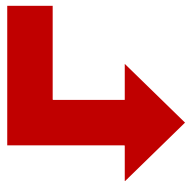
(22, 17)

	id	name	categories	referralId	hasPerk	location.lat	location.lng
0	4f8db193e4b05cdaa118f17b	Chinese	[{'id': '4bf58dd8d48988d145941735', 'name': 'C...	V-1578334510	False	51.616079	-0.132721
1	516d4b05e4b0aced3b68e637	Chinese Golden Buffet	[{'id': '4bf58dd8d48988d145941735', 'name': 'C...	V-1578334510	False	51.632861	-0.126294
2	4c5b0bfa15aaef3b4a19c580	Honeymoon Chinese Restaurant	[{'id': '4bf58dd8d48988d145941735', 'name': 'C...	V-1578334510	False	51.632137	-0.128725
3	4bd4a67f6f649521bdd36dec	Wood Green Chinese Restaurant	[{'id': '4bf58dd8d48988d145941735', 'name': 'C...	V-1578334510	False	51.614642	-0.128742
4	51f2d55a498e87240a32f1bd	Chinese Restaurant	[{'id': '4bf58dd8d48988d145941735', 'name': 'C...	V-1578334510	False	51.595484	-0.112953

- **Barnet = 22**
- Tower Hamlets = 50
- Southwark = 50

Conclusion

Barnet is the most suitable location for a new Chinese restaurant.

Reasoning:

- Highest Chinese population (8259)
- Fewest local competitors (22)
- Relatively cheap property price (£618,697)

Limitations:

- Boroughs in dataset limited by source (19/32)
- Property prices not guaranteed (average)
- Dataset for London populations outdated (2011)