

APPLIED DATA SCIENCE CAPSTONE

**PREDICTING THE MOST APPROPRIATE LONDON
NEIGHBORHOODS TO OPEN A MEDITERRANEAN FUSION
RESTAURANT**

BACKGROUND

- ▶ Mediterranean cuisine is getting a lot of traction as a diet which provides a plethora of health benefits
- ▶ Modern societies promote more and more a healthy and balanced way of life
- ▶ People seek more and more new experiences that stimulate their senses in unique and unprecedented ways
- ▶ This is more profound in western metropolitan cities

SCOPE

- ▶ Find the most suitable neighborhoods in London to open a new Mediterranean fusion restaurant that offers
 - ▶ modern cuisine with focus on quality and seasonal ingredients
 - ▶ unique gastronomic experience

DATA ACQUISITION AND PROCESSING

- ▶ Data sources
 - ▶ Wikipedia data regarding the locations in London
 - ▶ YouGov's report on the favourite cuisines in London based on people's profiles
 - ▶ Foursquare API to analyze the existing venues in the locations of interest

METHODOLOGY

- ▶ Initial data acquisition of the London areas
- ▶ Exclusion of areas based on demographics data
- ▶ Extraction of the existing venues in the locations of interest
 - ▶ We use the venue category and pricing in our analysis as both of these attributes have direct relation to the type of the restaurant we want to open.
- ▶ Initial exploration of a single location to confirm our assumptions
- ▶ Application of k-means clustering to all areas of interest to identify common patterns in the venues located in each one
- ▶ Definition of the final set of neighborhoods under consideration

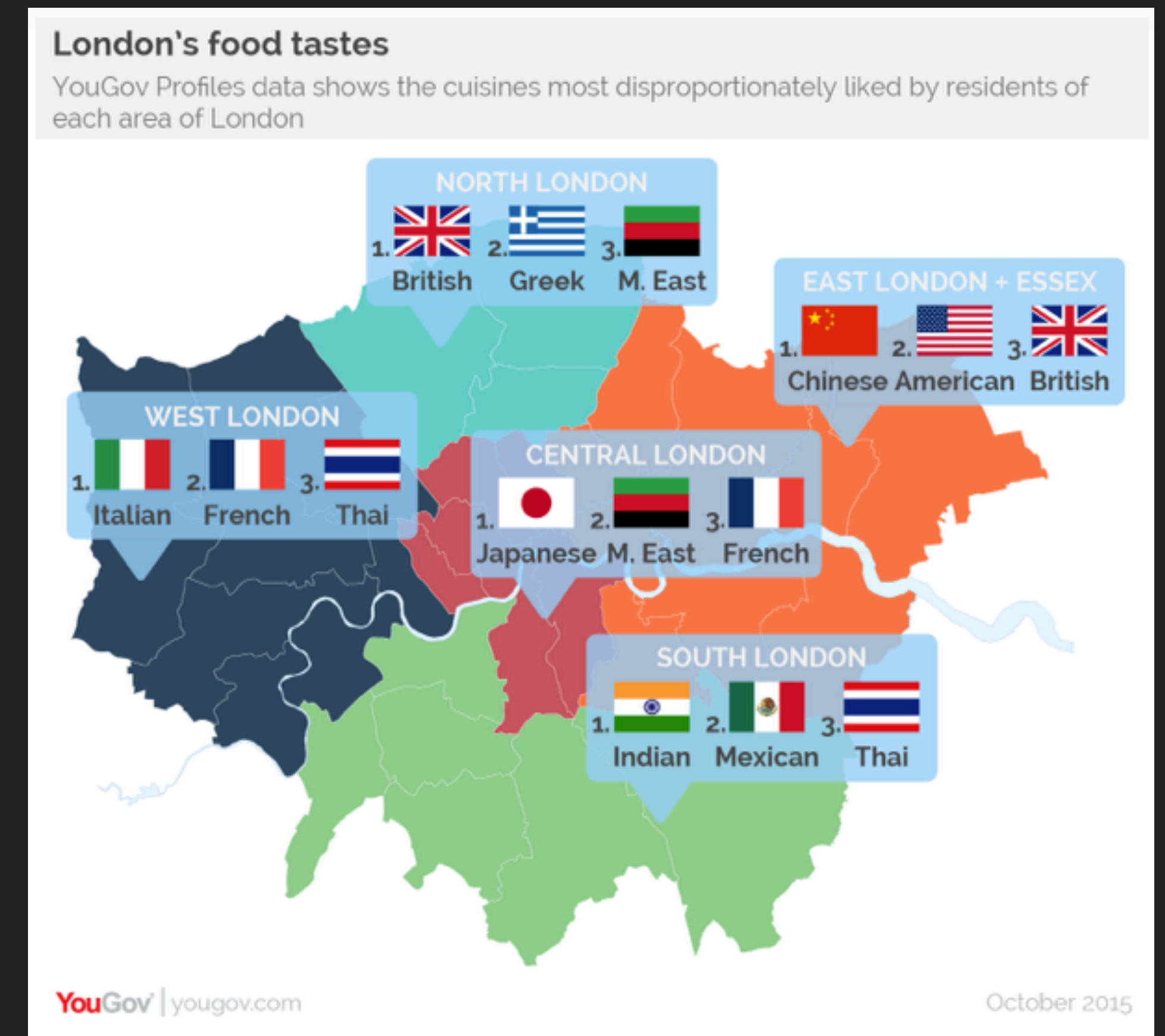
INITIAL DATAFRAME

- ▶ The initial dataframe extracted from Wikipedia contains
 - ▶ Neighborhood
 - ▶ Borough
 - ▶ Postcode

	Location	Borough	Postcode
0	Abbey Wood	Bexley, Greenwich	SE2
1	Acton	Ealing, Hammersmith, Fulham	W3
1	Acton	Ealing, Hammersmith, Fulham	W4
6	Aldgate	City	EC3
7	Aldwych	Westminster	WC2

EXCLUSION OF NEIGHBORHOODS BASED ON DEMOGRAPHICS

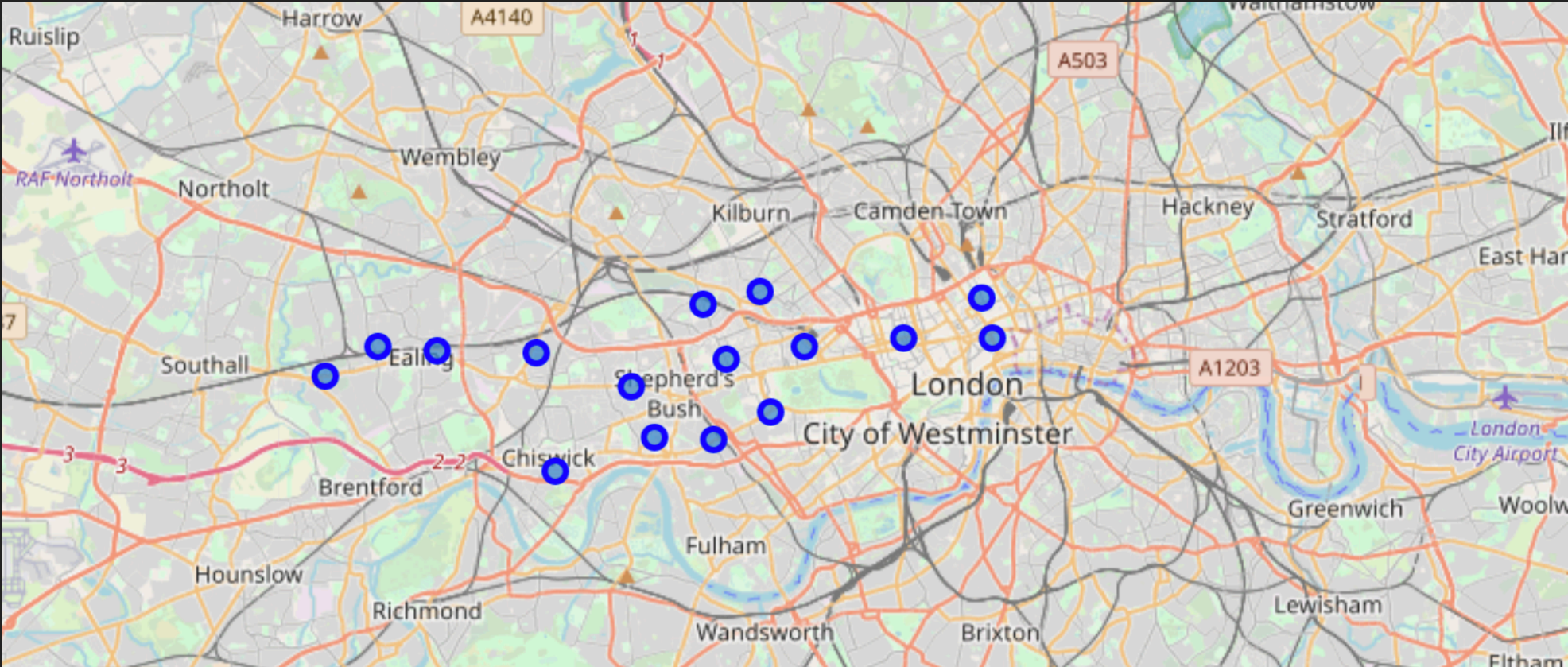
- ▶ Based on the YouGov's report, we keep only the locations in West London.
- ▶ People in West London mostly prefer Mediterranean based cuisines (French, Italian)
- ▶ We also know that West London is one of the most expensive locations in London and the majority of its habitants have high income



Source: YouGov

<https://yougov.co.uk/topics/politics/articles-reports/2015/10/13/mapped-londons-food-tastes>

FINAL LIST OF NEIGHBORHOODS UNDER CONSIDERATION



	Postcode	Borough	Location	Latitude	Longitude
0	W1	Westminster, Camden	Mayfair, Chinatown, Soho, Marylebone (also St ...	51.516479	-0.14816
1	W10	Kensington, Chelsea	North Kensington	51.523289	-0.21343
2	W11	Kensington, Chelsea	Notting Hill	51.512249	-0.20626
3	W12	Fulham, Hammersmith	White City, Shepherd's Bush, Wormwood Scrubs	51.506451	-0.23691
4	W13	Ealing	West Ealing	51.514530	-0.31951
5	W14	Fulham, Hammersmith	West Kensington	51.495739	-0.20980
6	W2	Westminster	Bayswater, Paddington	51.514938	-0.18049
7	W3	Fulham, Hammersmith, Ealing	Acton	51.513241	-0.26746
8	W4	Fulham, Hammersmith, Ealing, Hounslow	Gunnersbury, Chiswick, Bedford Park, Grove Park	51.489441	-0.26194
9	W5	Ealing	Ealing	51.514069	-0.29991
10	W6	Fulham, Hammersmith	Hammersmith	51.496151	-0.22942
11	W7	Ealing	Hanwell	51.508781	-0.33630
12	W8	Kensington, Chelsea	Holland Park	51.501610	-0.19175
13	W9	Westminster	Maida Vale, Little Venice	51.525871	-0.19526
14	WC1	Islington, Camden	Holborn, Bloomsbury, St Pancras, King's Cross	51.524502	-0.12273
15	WC2	Westminster, Camden	St Giles, Covent Garden, Aldwych, Charing Cross	51.516510	-0.11967

VENUES EXPLORATION USING FOURSQUARE API

- ▶ For each area, we search for the top 100 venues with pricing tiers 3 and 4 (the most expensive) within a radius of 500m
- ▶ We create a dataframe with their category and we group by it to find the 10 most common venue categories per area

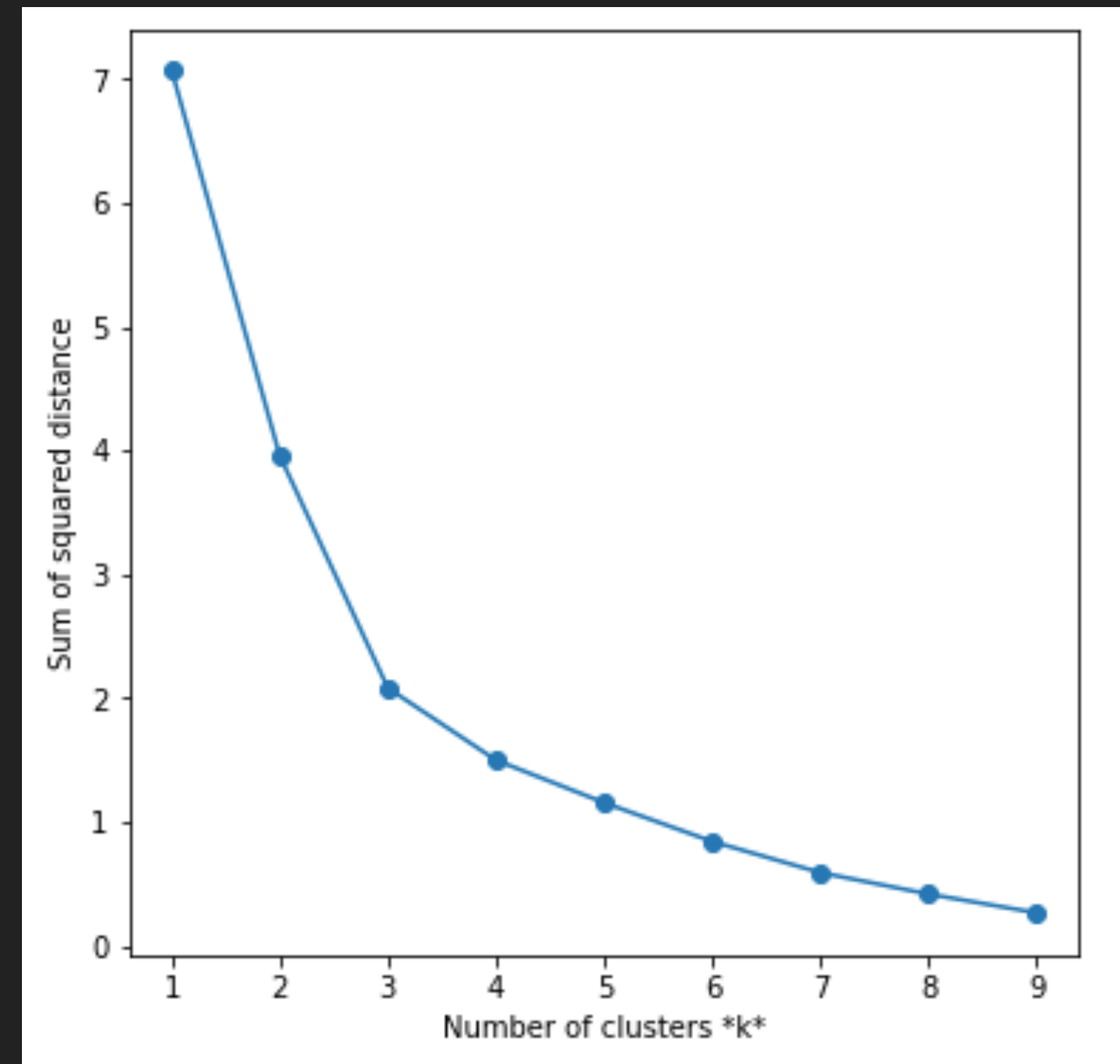
	venue	freq
0	Hotel Bar	0.22
1	Beer Bar	0.11
2	Tapas Restaurant	0.11
3	Seafood Restaurant	0.11
4	Restaurant	0.11
5	Turkish Restaurant	0.11
6	Brasserie	0.11
7	Cocktail Bar	0.11
8	Cycle Studio	0.00
9	Movie Theater	0.00

	Location	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	Aldwych	Hotel Bar	Turkish Restaurant	Beer Bar	Brasserie	Seafood Restaurant	Restaurant	Cocktail Bar	Tapas Restaurant	Gourmet Shop	Café
1	Bayswater	Steakhouse	Restaurant	French Restaurant	Seafood Restaurant	Cocktail Bar	Dim Sum Restaurant	Hotel Bar	Turkish Restaurant	Gourmet Shop	Beer Bar
2	Bedford Park	French Restaurant	Turkish Restaurant	Tapas Restaurant	Beer Bar	Brasserie	Café	Chinese Restaurant	Cocktail Bar	Cycle Studio	Dim Sum Restaurant
3	Bloomsbury	Szechuan Restaurant	Cocktail Bar	Hotel Bar	Tapas Restaurant	Turkish Restaurant	Gourmet Shop	Beer Bar	Brasserie	Café	Chinese Restaurant
4	Charing Cross	Hotel Bar	Turkish Restaurant	Beer Bar	Brasserie	Seafood Restaurant	Restaurant	Cocktail Bar	Tapas Restaurant	Gourmet Shop	Café

- ▶ The above figure presents the top 10 most common venue categories at a given location

K-MEANS CLUSTERING

- ▶ In order to define the optimal k index for the k-means algorithm, the Elbow method is used



RESULTS

- ▶ From the individual dataframes per cluster, we can identify what are the common attributes of the venues in the cluster.
 - ▶ In cluster 1, the most common venue is Gastrohub followed by Turkish restaurant as the second most common venue. Locations in cluster 1 are not the most suitable as the most common venues serve food.
 - ▶ In cluster 2, the three most common venues are French, Turkish, and Tapas restaurants. Again, these are not the locations that we would consider.
 - ▶ In cluster 3, the most common venue is Lounge. However, the second most common is Turkish restaurant. So, we may consider locations in this cluster, but further analysis is required to closely evaluate the competition and decide to proceed or not.
 - ▶ In cluster 4, the most common venue is either a hotel bar or Asian restaurant. There are also various types of venues as second most common ones. This makes this cluster the most interesting of all, since there is space for the new restaurant we are considering. The locations in this cluster are the first we would further explore in order to reach a final decision.

CONCLUSION

- ▶ Using data from a variety of sources, we evaluated the neighborhoods of West London in order to find the most appropriate ones in order to open a Mediterranean fusion restaurant.
- ▶ We concluded to the subset of areas that seems the most suitable. However, further analysis, both quantitative and qualitative, is required in order to limit the number of potential areas of interest.
- ▶ Such analysis would include computations regarding distances among competing restaurants, distances from means of public transport, distances from suppliers, etc.
- ▶ Also, the cost of lease and the availability of spaces to open a restaurant are important factors.
- ▶ Finally, qualitative metrics such as customer satisfaction from existing restaurants, customer perception of each neighborhood, etc. would also affect the final decision.