A NEW RESTAURANT IN LONDON

FINAL PROJECT OF COURSERA CAPSTONE

TABLE OF CONTENT

1. *DESCRIPTION OF THE PROBLEM AND A DISCUSSION OF THE BACKGROUND*

- Let's suppose, an investor has enough time and money, as well as a passion to open the best eating spot in London. What type of restaurant would it be? What would be the best place for it? Is there a better way to answer these questions rather than guessing? What if there is a way to cluster city neighborhoods, based on their restaurant similarity? What if we can visualize these clusters on a map? What if we might find what type of restaurant is the most and least popular in each location?
- ***Target audience:***

  Investors, entrepreneurs, and chefs interested in opening a restaurant in London, who may need a piece of objective advice of what type of restaurant would be more successful and where exactly it should be opened.

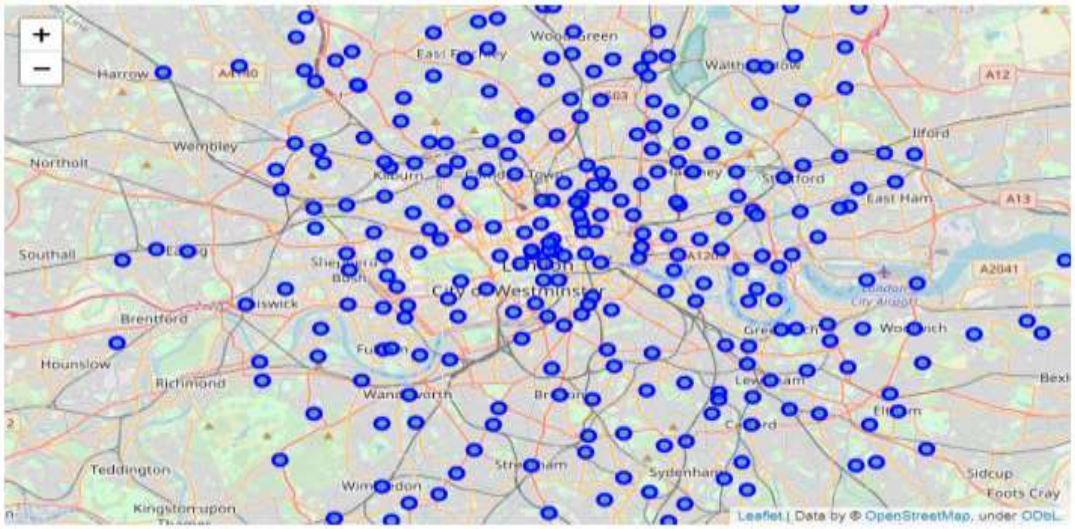
2. *DESCRIPTION OF THE DATA AND HOW IT WILL BE USED TO SOLVE THE PROBLEM*

- Step 1. Using a table on https://en.wikipedia.org/wiki/List_of_areas_of_London, collect information about London boroughs and locations, excluding records whose "Post Town" is not London.
- Step 2. Use the Geopy and Folium library to get the coordinates of every locations and map geospatial data on a London map.
- Step 3. Using Foursquare API, collect the top 100 restaurants and their categories for each location within a radius 500 meters.
- Step 4. Group collected restaurants by location and by taking the mean of the frequency of occurrence of each type, preparing them for clustering.
- Step 5. Cluster restaurants by k-means algorithm and analyze the top 10 most common restaurants in each cluster.
- Step 6. Visualize clusters on the map, thus showing the best locations for opening the chosen restaurant.

3.METHODOLOGY

- Section 1 :Using Wikipedia we collected all of London's neighborhoods
- Section 2 : Using Geospatial libraries we added geographical coordinates
- With some cleaning and wrangling we obtained 288 locations.

Out[23]:



4.EXPLORING LONDON RESTAURANTS

- We utilize the Foursquare API to get the top 100 restaurants in each neighborhood.
- Thus we obtained 7801 restaurants of 130 individual types.
- We applied one-hot encoding and grouped them by taking the mean of the frequency of occurrence for each type.
- Then we clustered restaurants using the k-means algorithm based on their similarity. The k-means is an unsupervised machine learning algorithm for clustering unlabeled data.

5.RESULTS

- Cluster 1:

	Borough	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
count	134	134.0	134	134	134	134	134	134	134	134	134	134
unique	33	NaN	20	27	36	37	49	46	50	49	45	52
top	Tower Hamlets	NaN	Café	Café	Restaurant	Café	Italian Restaurant	Fast Food Restaurant	Falafel Restaurant	Ethiopian Restaurant	English Restaurant	Dumpling Restaurant

- The most common restaurant: Café
- The 10th most common restaurant: Dumpling Restaurant

5.RESULTS

- Cluster 2:

	Borough	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
count	30	30.0	30	30	30	30	30	30	30	30	30	30
unique	17	NaN	8	16	19	21	18	19	21	19	18	17
top	Barnet	NaN	Café	Fast Food Restaurant	Bakery	Pizza Place	Fast Food Restaurant	Diner	Burger Joint	Falafel Restaurant	Dumpling Restaurant	Eastern European Restaurant

- The most common restaurant: Café
- The 10th most common restaurant: Eastern European Restaurant

5.RESULTS

- Cluster 3:

	Borough	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
count	1	1.0	1	1	1	1	1	1	1	1	1	1
unique	1	NaN	1	1	1	1	1	1	1	1	1	1
top	Croydon	NaN	English Restaurant	Yoshoku Restaurant	Currywurst Joint	Deli / Bodega	Dim Sum Restaurant	Diner	Doner Restaurant	Donut Shop	Dumpling Restaurant	Eastern European Restaurant

- The most common restaurant: English Restaurant
- The 10th most common restaurant: Eastern European Restaurant

5.RESULTS

- Cluster 4:

	Borough	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
count	7	7.0	7	7	7	7	7	7	7	7	7	7
unique	6	NaN	2	5	7	5	5	5	5	5	5	4
top	Newham	NaN	Indian Restaurant	Bakery	Mediterranean Restaurant	Bakery	Eastern European Restaurant	Falafel Restaurant	Ethiopian Restaurant	English Restaurant	Empanada Restaurant	Dumpling Restaurant

- The most common restaurant: Indian Restaurant
- The 10th most common restaurant: Dumpling Restaurant

5.RESULTS

- Cluster 5:

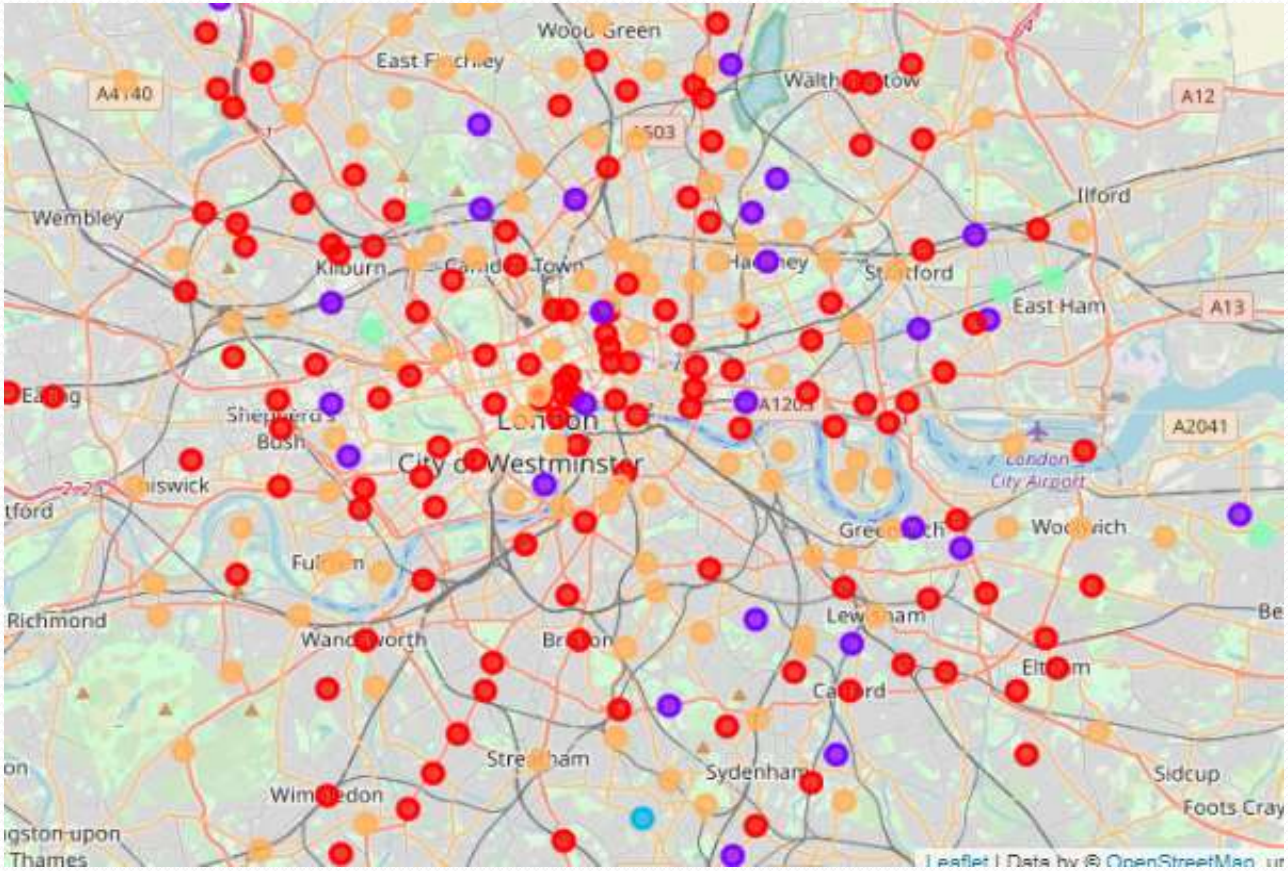
	Borough	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
count	112	112.0	112	112	112	112	112	112	112	112	112	112
unique	29	NaN	16	27	32	40	36	42	39	46	44	44
top	Barnet	NaN	Café	Café	Italian Restaurant	Italian Restaurant	Sandwich Place	Fast Food Restaurant	Falafel Restaurant	Falafel Restaurant	English Restaurant	English Restaurant

- The most common restaurant: Café
- The 10th most common restaurant: English Restaurant

5.RESULTS

- Visualizing Clusters:

Cluster 1 - red dots
Cluster 2 - purple dots
Cluster 3 - blue dot
Cluster 4 - green dots
Cluster 5 - orange dots



6.DISCUSSION

- Analyzing the most popular restaurants in each cluster the stakeholder should select the least popular types within the top10 as a safe choice. In our recommendations we advise selecting from 9th or 10th positions. This selection is a reasonable balance between starting an ordinary business and having no costumers.
- Recommendations, based on description of each cluster:
  Cluster 1 Locations: Empanada or English Restaurant
  Cluster 2 Locations: Dumpling or Eastern European Restaurant
  Cluster 3 Locations: Dumpling or Eastern European Restaurant
  Cluster 4 Locations: Empanada or Dumpling Restaurant
  Cluster 5 Locations: English or Ethiopian Restaurant
- After the type of restaurant is chosen, it is time to select a right place. Using the map of clusters created in Part 5 the solution is obvious.

7.CONCLUSSION

- In this report we established a methodology to determine what the most promising type of restaurant is, and where it should be opened.
- This type of analysis can be applied to any city of your choice that has available geospatial information.
- This type of analysis can be applied to any type of venue(shopping, clubs, pubs, etc) that is available in Foursquare database.