

# Capstone Project - Open a Chinese restaurant in North London

# Introduction

- ▶ Huge amount of Chinese people currently residing in London
- ▶ Numerous Chinese students and tourists come to London every year
- ▶ Chinese restaurants are always popular for both Chinese people and those who are not Chinese but love Chinese food
- ▶ There are not many fine Chinese restaurants in North London
- ▶ The number of local Chinese people and students/tourists in North area is massive
- ▶ It would be a wonderful idea to open a fine Chinese restaurant in North London

# Problem - what location should we choose to open a Chinese restaurant in North London?

- ▶ The stakeholders and audiences are those who are interested in opening a Chinese restaurant in North London so that it could meet the local needs
- ▶ The stakeholders and audiences would care about this problem because a suitable location is extremely important for the success of the Chinese restaurant

# Data 1

- ▶ 'List of Areas of London' Data (containing location, borough, postcode, and etc.) which is similar to the datasets used in the previous weeks of this course
- ▶ Available on: [https://en.wikipedia.org/wiki/List\\_of\\_areas\\_of\\_London](https://en.wikipedia.org/wiki/List_of_areas_of_London)
- ▶ Will be scrapped from the Wikipedia page, wrangled, cleaned and then read into a pandas dataframe so that it is a structured format which can be ready to use
- ▶ For example, the data will be reduced to North London Areas and only relevant features will be selected for further analysis

# Data 2

- ▶ 'Geocoder' package will be used to obtain the latitude and longitude of the locations
- ▶ 'Foursquare API' will be used to obtain location data which will be helpful for searching and exploring various neighborhoods information such as surrounding venues (especially restaurants) in the North London area

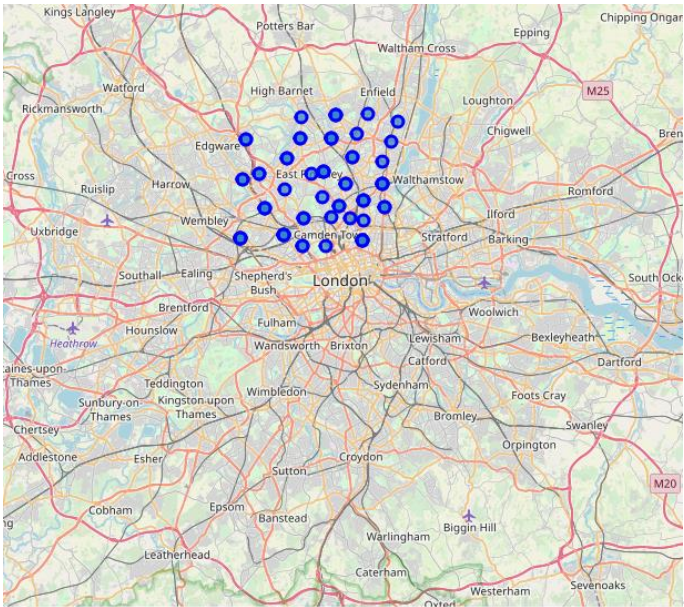
# Methodology

- ▶ Data collection and preparation
  - ▶ Scrape the table from the Wikipedia page
  - ▶ Wrangle and clean the data
- ▶ Get the latitude and longitude of each location
  - ▶ Define a function to obtain latitude and longitude
  - ▶ Obtain the data of latitude and longitude and append them to the dataframe
- ▶ Explore the locations in North London
  - ▶ Create a map of London with North London Locations superimposed on top
  - ▶ Explore the single Location Angel in the dataframe and get the top 100 venues
  - ▶ Explore Locations in North London
- ▶ Cluster Locations (using K-Means Clustering,  $k=5$ )
  - ▶ Examine Each Cluster

# Results - Initialization

	Postcode	Borough	Location	Latitude	Longitude
0	N1	Hackney	Hackney	51.53791	-0.09984
1	N1	Islington	Angel	51.53791	-0.09984
2	N10	Barnet	Colney Hatch	51.59107	-0.14852
3	N12	Barnet	Finchley	51.61592	-0.17674
4	N14	Enfield	Arnos Grove	51.63429	-0.13366

Structured Data



North London Locations

# Results - single Location Angel (top 100 venues within a radius of 1000 meters)

	name	categories	lat	lng
0	69 Colebrooke Row	Cocktail Bar	51.536480	-0.101207
1	Steve Hatt	Fish Market	51.538588	-0.099041
2	The Life Centre Islington	Yoga Studio	51.538056	-0.099171
3	Cass Art	Arts & Crafts Store	51.536379	-0.101013
4	The Bill Murray	Pub	51.536226	-0.098750

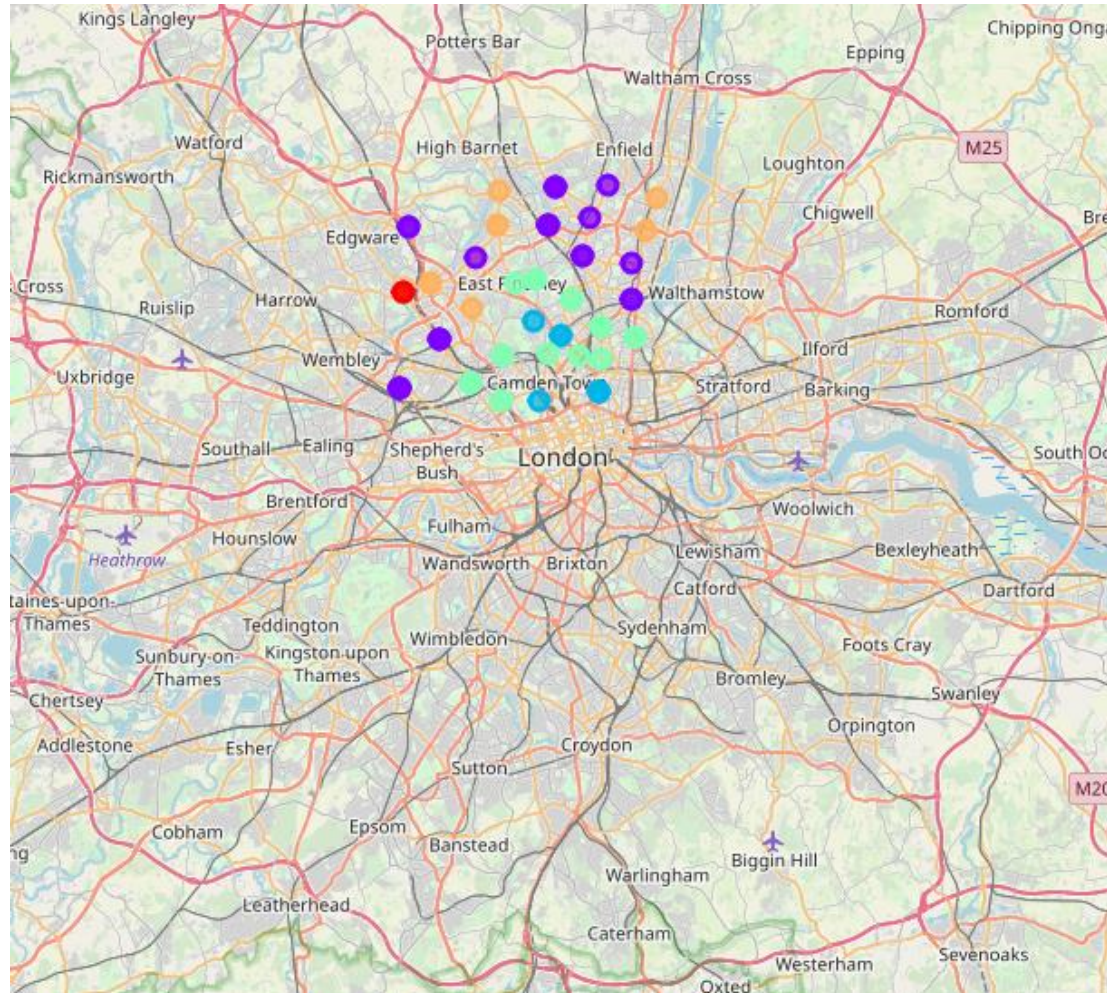
	count
Pub	11
Mediterranean Restaurant	4
Gastropub	4
Burger Joint	4
French Restaurant	4
Bakery	3
Cocktail Bar	3
Boutique	3
Japanese Restaurant	2
Furniture / Home Store	2



# Results - Locations in North London (top 100 venues within a radius of 1000 meters)

	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	Angel	Pub	French Restaurant	Mediterranean Restaurant	Burger Joint	Gastropub	Boutique	Cocktail Bar	Bakery	Arts & Crafts Store	Coffee Shop
1	Archway, Upper Holloway	Pub	Coffee Shop	Pizza Place	Italian Restaurant	Hotel	Gastropub	Café	Park	Eastern European Restaurant	Music Venue
2	Arkley	Grocery Store	Discount Store	Train Station	Café	Chinese Restaurant	Pizza Place	Fast Food Restaurant	Bookstore	Salon / Barbershop	Gym / Fitness Center
3	Arnos Grove	Grocery Store	Fast Food Restaurant	Gym / Fitness Center	Coffee Shop	Bus Stop	Pizza Place	Park	Electronics Store	Café	Pub
4	Barnet Gate, The Hale, Mill Hill	Grocery Store	Discount Store	Train Station	Café	Chinese Restaurant	Pizza Place	Fast Food Restaurant	Bookstore	Salon / Barbershop	Gym / Fitness Center

## Results - Cluster Locations (using K-Means Clustering, k=5)



# Results - Examine Each Cluster

- ▶ Details in the Notebook
- ▶ Will be summarized in the next slide

# Results - Summary

- ▶ Chinese restaurant is the top 5 most common venues in these locations: 'Brent Cross, Childs Hill', 'Cricklewood', 'Dollis Hill, Neasden', 'Hackney, Shacklewell, Stamford Hill', 'Stoke Newington, West Hackney'
- ▶ The differences between 5 Clusters are obvious in terms of the most common venues (in the Examine Clusters Section)
- ▶ Chinese restaurant is popular in Cluster 2, whereas French, Mediterranean and Italian restaurants are common in Cluster 3
- ▶ Cluster 4, compared to the other 4 Clusters, is the only Cluster that does not have many top restaurants. Instead, Café, Pub and Grocery Store are more popular

# Discussion and Conclusion

- ▶ Since Cluster 4 is the only Cluster that does not have many top restaurants, I would recommend that opening a new Chinese restaurant in Cluster 4 is the most suitable
- ▶ Chinese restaurant is the top 5 most common venues in the location of 'Dollis Hill, Neasden'. In addition, the location of 'West Hampstead' in Cluster 4 is the closest to 'Dollis Hill, Neasden'
- ▶ 'West Hampstead' could potentially be a wonderful location for a new Chinese restaurant