Where to open a Chinese restaurant in London?

Coursera Capstone Project

Content

- Data preparation
- Clustering methodology
- Conclusion

Data preparation – Source 1

Data of London boroughs: area size, population, geo-coordinates scrapped from Wikipedia

Web



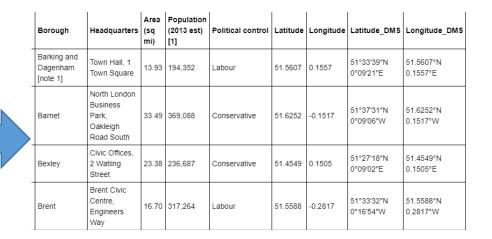


Table on Wikipedia

Pandas data frame

Data preparation – Source 2

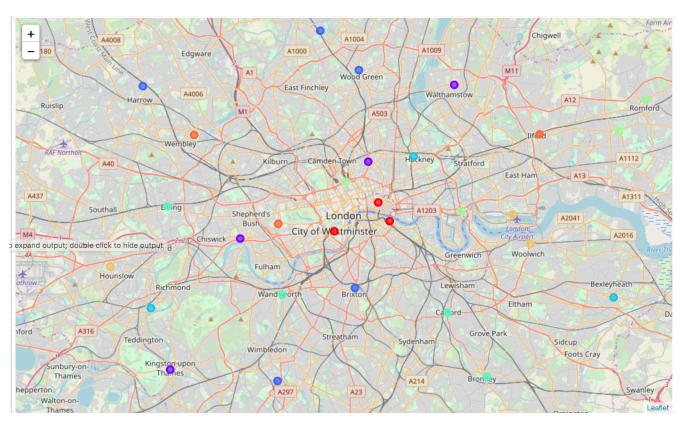
London venue data from Foursquare: venue name, geo-coordinates, venue type

	Borough	Latitude	Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Barking and Dagenham [note 1]	51.5607	0.1557	Central Park	51.559560	0.161981	Park
1	Barking and Dagenham [note 1]	51.5607	0.1557	The Eva Hart (Wetherspoon)	51.570460	0.130342	Pub
2	Barking and Dagenham [note 1]	51.5607	0.1557	The Range	51.575550	0.180254	Furniture / Home Store
3	Barking and Dagenham [note 1]	51.5607	0.1557	Costa Coffee	51.576890	0.179497	Coffee Shop
4	Barking and Dagenham [note 1]	51.5607	0.1557	Harrow Lodge Park	51.555648	0.197926	Park

Sample venue data by borough

Clustering

Run K-means to cluster boroughs into 8 clusters



Clustering – Results analysis

Below attributes are desirable for opening a new Chinese restaurant:

- Significant number of restaurants out of all venues
- Ratio of existing Chinese restaurants out of all restaurants is relatively low
- Large number of pubs
- Large number of hotels
- Relatively fewer gyms
- Densely populated
- Office presence

Conclusion

Cluster 0 has the most desirable attributes:

- Relatively high number of restaurants
- No Chinese restaurants
- Plenty of pubs and hotels
- Fewer gyms

Out of the 3 preferred boroughs in Cluster 0, Southwark is the most preferred because it is more densely populated than the other two boroughs.