

London city tours by venue themes

Danuphan Suwanwong
June 5, 2019



Agenda

1. Motivation
2. Data Acquisition
3. Analysis
4. Discussion & Conclusion





Motivation





Motivation

A London tour agency want to create new city tours based on venue theme.



Data Acquisition



London boroughs

1. City of London*
2. City of Westminster
3. Kensington and Chelsea
4. Hammersmith and Fulham
5. Wandsworth
6. Lambeth
7. Southwark
8. Tower Hamlets
9. Hackney
10. Islington
11. Camden
12. Brent
13. Ealing
14. Hounslow
15. Richmond
16. Kingston
17. Merton



18. Sutton
19. Croydon
20. Bromley
21. Lewisham
22. Greenwich
23. Bexley
24. Havering
25. Barking and Dagenham
26. Redbridge
27. Newham
28. Waltham Forest
29. Haringey
30. Enfield
31. Barnet
32. Harrow
33. Hillingdon

*The City of London is not a borough but an independent county.

Source: https://en.wikipedia.org/wiki/Greater_London



Venues by boroughs

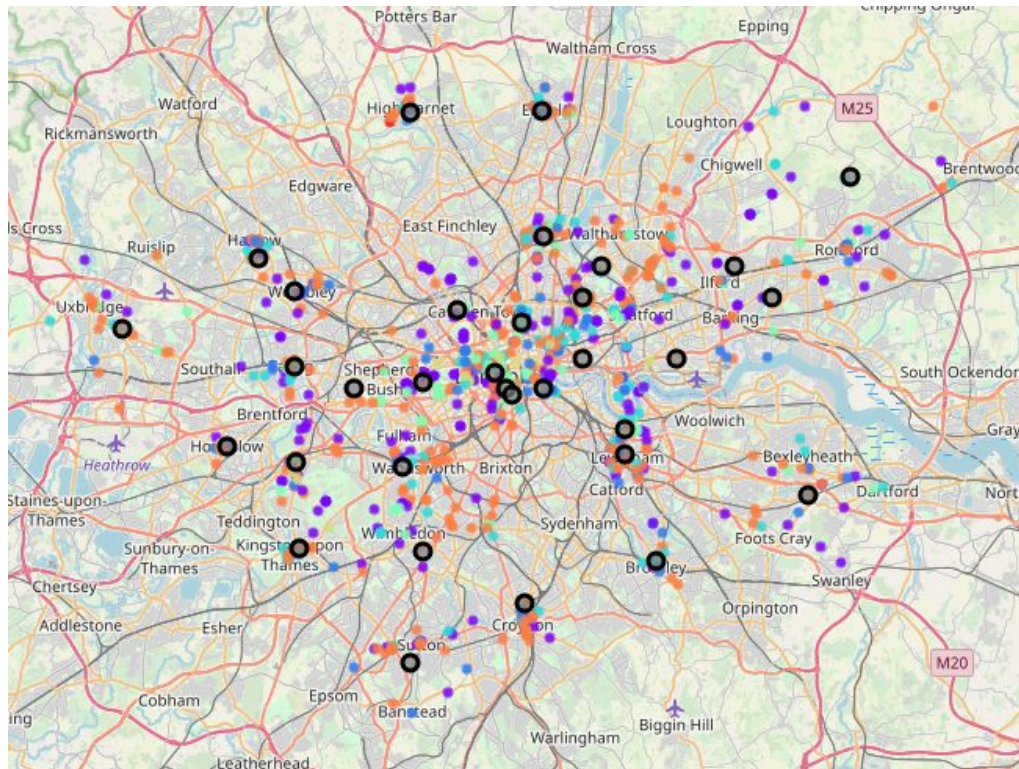
- Venues are from the [“Get Venue Recommendations” endpoint](#) combining with the ["Get Venue Categories" endpoint](#) for venue category tree provided by [Foursquare API](#).
- To get a high overview of venues, we consider their root categories rather than the original category returned by the API since original categories of each venue can be **very sparse**.

Original categories	Root categories
Dim Sum Restaurant	Food
Movie Theater	Arts & Entertainment



Analysis





Map of London boroughs (black dots) and recommended venues (colored dot)



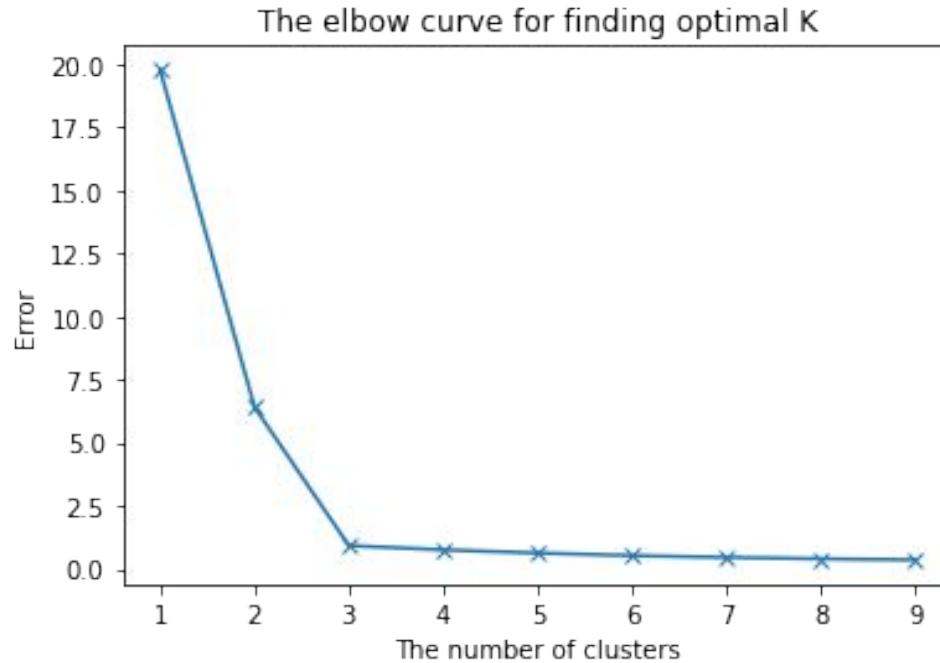
Structure data

- Apply **one-hot encoding** to venue categories
- **Group venues** by their boroughs and compute proportion of venue categories per borough

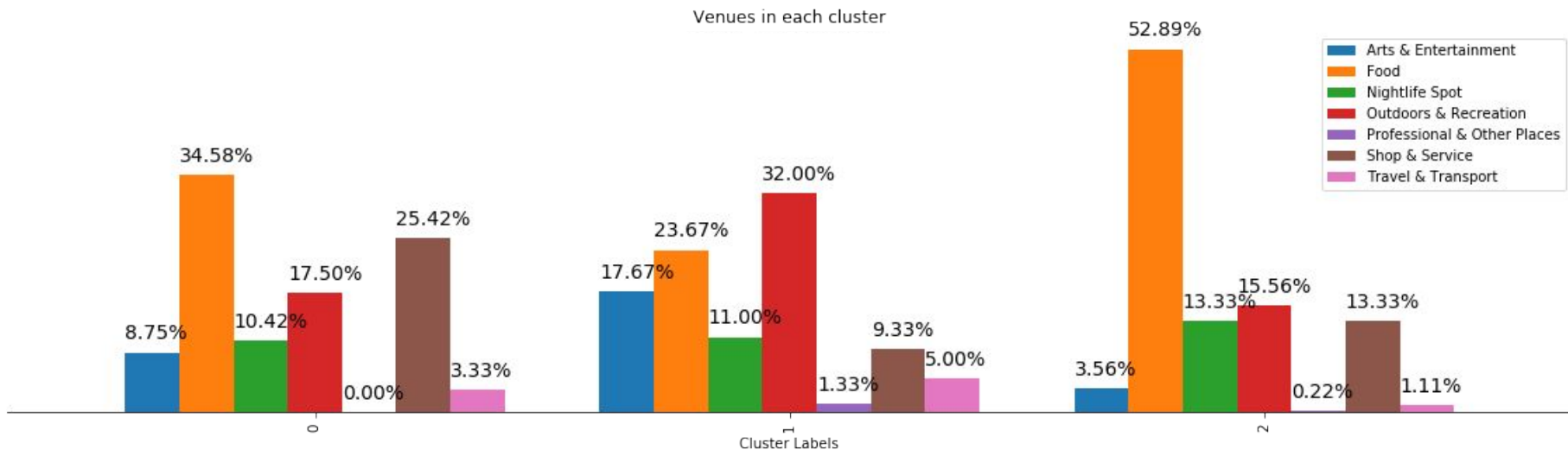
	Borough	Arts & Entertainment	Food	Nightlife Spot	Outdoors & Recreation	Professional & Other Places	Shop & Service	Travel & Transport
22	Lambeth	0.200000	0.333333	0.133333	0.233333	0.033333	0.000000	0.066667
6	City of London	0.033333	0.066667	0.100000	0.633333	0.000000	0.133333	0.033333
5	Camden	0.200000	0.200000	0.100000	0.333333	0.000000	0.100000	0.066667
18	Hounslow	0.000000	0.466667	0.033333	0.066667	0.000000	0.366667	0.066667
15	Harrow	0.000000	0.533333	0.200000	0.100000	0.000000	0.166667	0.000000



K-Mean: find optimal K clusters



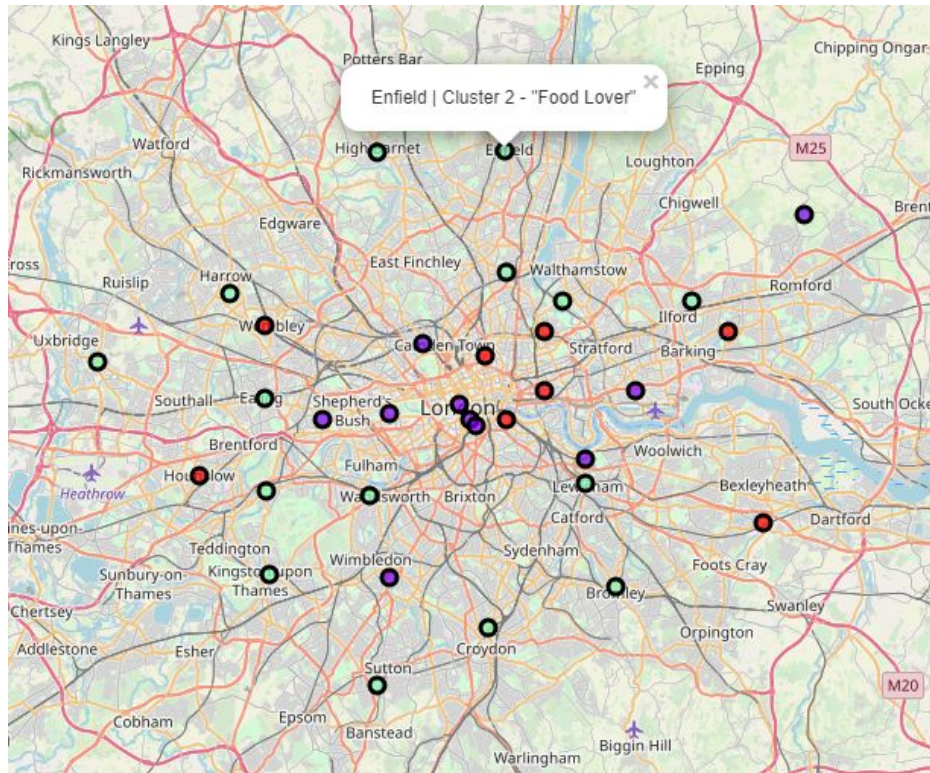
K-Mean: Results



“Food and Shopping Lover”

“Outdoors and Entertainments Lover”

“Food Lover”



Map of London boroughs with assigned cluster labels.

Labels: “Food and Shopping Lover” (red), “Outdoors and Entertainments Lover” (blue), “Food Lover” (green)



Discussion & Conclusion





Discussion & Conclusion

- Not surprisingly, the **central of London has a good balance of venues** comparing to the outer boroughs which mainly have a popularity of food. So not recommend outer boroughs for a tour as tourists would not come for only food.
- Venues from Foursquare API may not be restricted to given boroughs, so **some venues can overlap** other areas
- Possible expansion to other cities, and then analysis individually or even simultaneously.
- K-Mean algorithm gives a **meaningful clustering** (city themes)