

A guide for new-to-London students

By Louise Huang, 23rd March, 2021

Introduction:

This project aims to provide a guide for new-to-London college students, especially international students, to explore better facilities around their campus. This guide could help them make smart and efficient decisions on deciding neighborhood out of all the districts in London, UK. The first part of the guide is specifically for UCL/LSE/KCL students as their campus is located in central London. By utilizing this guide, students could find best area to try exotic cuisine, find home food or seek that niche bookshop area full of cute cafes.

Problem which tried to solve:

1. List and visualize Central London district with specific facilities
2. Which is the best neighborhood for you to find your home food when you get homesick?
3. Where would you go for an experience of theatre-going and book-shopping, aka art and cultural stuff?

Data used in this guide:

Foursquare API is used as the primary data source for this project. As a database of millions of places around the world, this project utilized Foursquare places API to provides location search, sharing and details.

List of London District Name and Post code is scrapped from Wikipedia. Latitude and the longitude coordinates of each district come from Geocoder package in

Python.

After data cleaning and wrangling, I performed K-means clustering algorithm to segment neighborhoods for better understanding and solving our problems mentioned above.

Methodology:

- Data retrieval, exploration and wrangling
- Performing K-means clustering algorithm to segment neighborhoods

Data retrieval, exploration and wrangling

List of postal code and district name is scrapped from Wikipedia. Latitude and Longitude of each neighborhood is retrieved using Geocoder from Geopy Library of Python.

Here is an overview of the database:

	Postcode district	Coverage	Local Authority Areas	Latitude	Longitude
0	WC1A	New Oxford Street	Camden	51.517165	-0.126810
1	WC1B	Bloomsbury, British Museum, Southampton Row	Camden	51.519140	-0.127759
2	WC1E	University College London, SOAS	Camden	51.522520	-0.133033
3	WC1H	St Pancras, UCL Institute of Education	Camden	51.524755	-0.131389
4	WC1N	Russell Square, Great Ormond Street	Camden	51.524460	-0.123254
5	WC1R	Gray's Inn	Camden	51.519375	-0.117610
6	WC1V	High Holborn	Camden, City of London	51.518660	-0.112102
7	WC1X	Kings Cross, Finsbury (west), Clerkenwell (north)	Camden, Islington	51.522049	-0.113885
8	WC2A	Lincoln's Inn Fields, Royal Courts of Justice,...	Camden, Westminster, City of London	51.516615	-0.116091
9	WC2B	Drury Lane, Kingsway, Aldwych	Camden, Westminster	51.514800	-0.121158
10	WC2E	Covent Garden	Westminster	51.511185	-0.121374
11	WC2H	Leicester Square, St. Giles	Camden, Westminster	51.515120	-0.128531
12	WC2N	Charing Cross	Westminster	51.509916	-0.125925
13	WC2R	Somerset House, Temple (west)	Westminster	51.510080	-0.121348
14	EC1A	St Bartholomew's Hospital	City of London, Islington	51.516355	-0.099137
15	EC1M	Clerkenwell, Farringdon	Islington, Camden, City of London	51.521011	-0.106675
16	EC1N	Hatton Garden	Camden, City of London	51.520027	-0.110511
17	EC1R	Finsbury, Finsbury Estate (west)	Islington, Camden	51.522350	-0.110057
18	EC1V	Finsbury (east), Moorfields Eye Hospital	Islington, Hackney	51.525715	-0.101704
19	EC1Y	St Luke's, Bunhill Fields	Islington, City of London	51.525630	-0.086289
20	EC2A	Shoreditch	Islington, Hackney, City of London	51.524365	-0.078885
21	EC2M	Broadgate, Liverpool Street	City of London, Tower Hamlets	51.517186	-0.087849
22	EC2N	Old Broad Street, Tower 42	City of London	51.516517	-0.084515
23	EC2R	Bank of England	City of London	51.516675	-0.089874
24	EC2V	Guildhall	City of London	51.517504	-0.094655
25	EC2Y	Barbican	City of London	51.517510	-0.093730
26	EC3A	St Mary Axe, Aldgate	City of London	51.513145	-0.078733
27	EC3M	Lloyd's of London, Fenchurch Street	City of London	51.511875	-0.079680
28	EC3N	Tower Hill, Tower of London	Tower Hamlets, City of London	51.508825	-0.076407
29	EC3R	Monument, Billingsgate	City of London	51.509559	-0.087146
30	EC3V	Cornhill, Gracechurch Street, Lombard Street	City of London	51.511420	-0.085105
31	EC4A	Fetter Lane	City of London, Westminster	51.515780	-0.110003
32	EC4M	St Paul's	City of London	51.512605	-0.094429
33	EC4N	Mansion House	City of London	51.512395	-0.090770
34	EC4R	Cannon Street	City of London	51.508651	-0.087484
35	EC4V	Blackfriars	City of London	51.510965	-0.094244
36	EC4Y	Temple	City of London, Westminster	51.511695	-0.106415
37	W1B	Portland Place, Regent Street	Westminster	51.520804	-0.145370
38	W1C	Oxford Street (west)	Westminster	51.514925	-0.148081
39	W1D	Soho (south east); Chinatown, Soho Square	Westminster	51.516370	-0.133809
40	W1F	Soho (north west)	Westminster	51.513095	-0.134890
41	W1G	Harley Street	Westminster	51.516065	-0.144746
42	W1H	Marylebone	Westminster	51.520290	-0.161740
43	W1J	Mayfair (south), Piccadilly	Westminster	51.508380	-0.138806
44	W1K	Mayfair (north), Grosvenor Square	Westminster	51.506095	-0.156574
45	W1S	Mayfair (east), Hanover Square, Seville Row, Ro...	Westminster	51.513575	-0.144274
46	W1T	Fitzrovia, Tottenham Court Road	Camden	51.517735	-0.133599
47	W1U	Marylebone	Westminster	51.515000	-0.151919
48	W1W	Great Portland Street, Fitzrovia	Westminster	51.522400	-0.143247
49	N1	Northern head district: Barnsbury (part), Cano...	Hackney, Islington, Camden	51.529690	-0.086970
50	N1C	Kings Cross Central	Camden	51.538120	-0.126721
51	SW1A	Whitehall, Buckingham Palace, the Houses of Pa...	Westminster	51.504456	-0.140785
52	SW1E	Northern 'Victoria' around Buckingham Gate (a ...	Westminster	51.498400	-0.142026
53	SW1H	Surrounds of St James's Park tube station cent...	Westminster	51.498610	-0.134037
54	SW1P	A projection around Parliament Square. Nearby ...	Westminster	51.495030	-0.138313
55	SW1V	between Vauxhall Bridge, Grosvenor (or Victori...	Westminster	51.494219	-0.142810
56	SW1W	South or lower Belgravia and Chelsea (a corner...	Westminster, Kensington and Chelsea	51.488545	-0.157410
57	SW1X	Rest of Belgravia (north of Eaton Square), Kni...	Westminster, Kensington and Chelsea	51.500645	-0.150956
58	SW1Y	All but the west of St James's	Westminster	51.507615	-0.132744
59	NW1	North Western head district: Marylebone (part)...	Camden, Westminster	51.533540	-0.146060

I used folium library to visualize the central London area:



Foursquare API is used to explore types of venues in each area and 10 top level categories are identified:

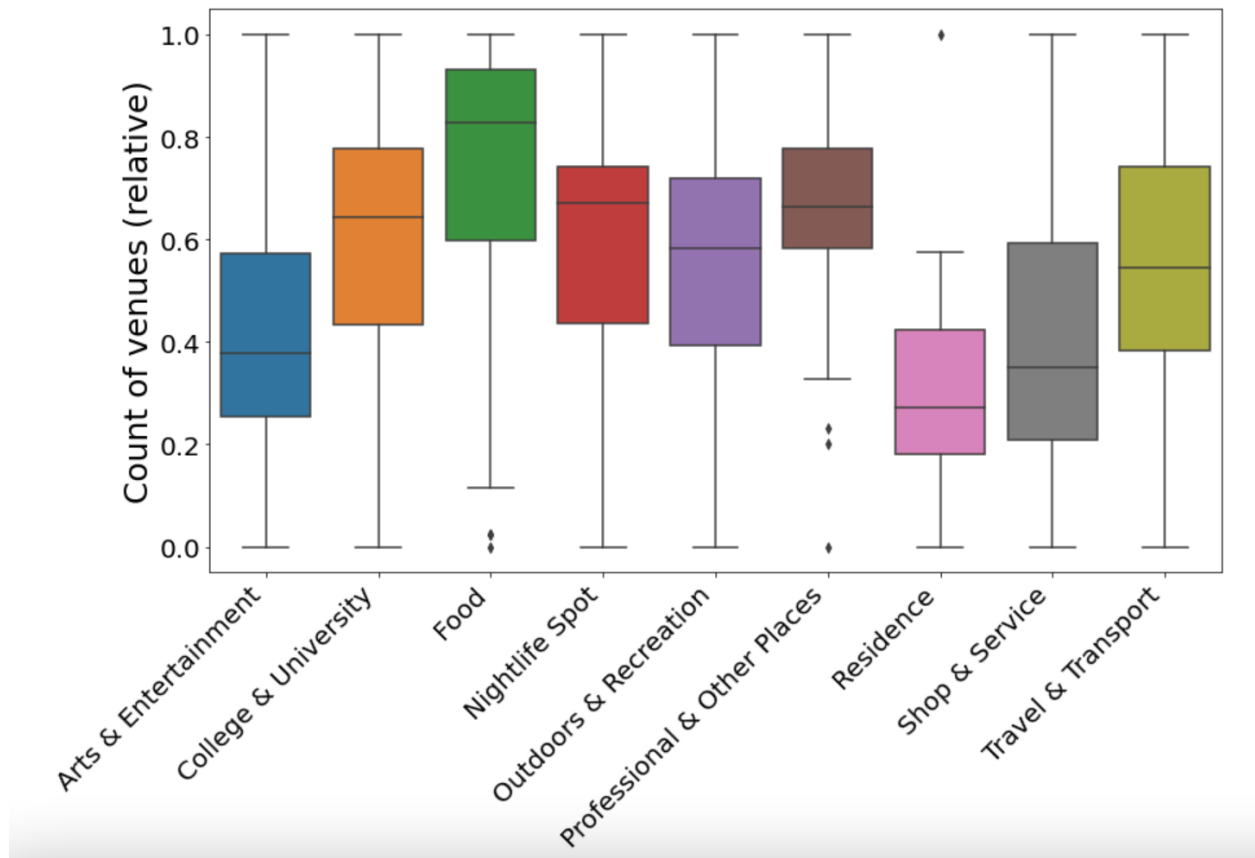
- Arts & Entertainment
- College & University
- Food
- Nightlife Spot
- Outdoors & Recreation
- Professional & Other Places
- Residence
- Shop & Service

- Travel & Transport

Here is a snapshot of part of the result data. We added up the total number of venues in each category.

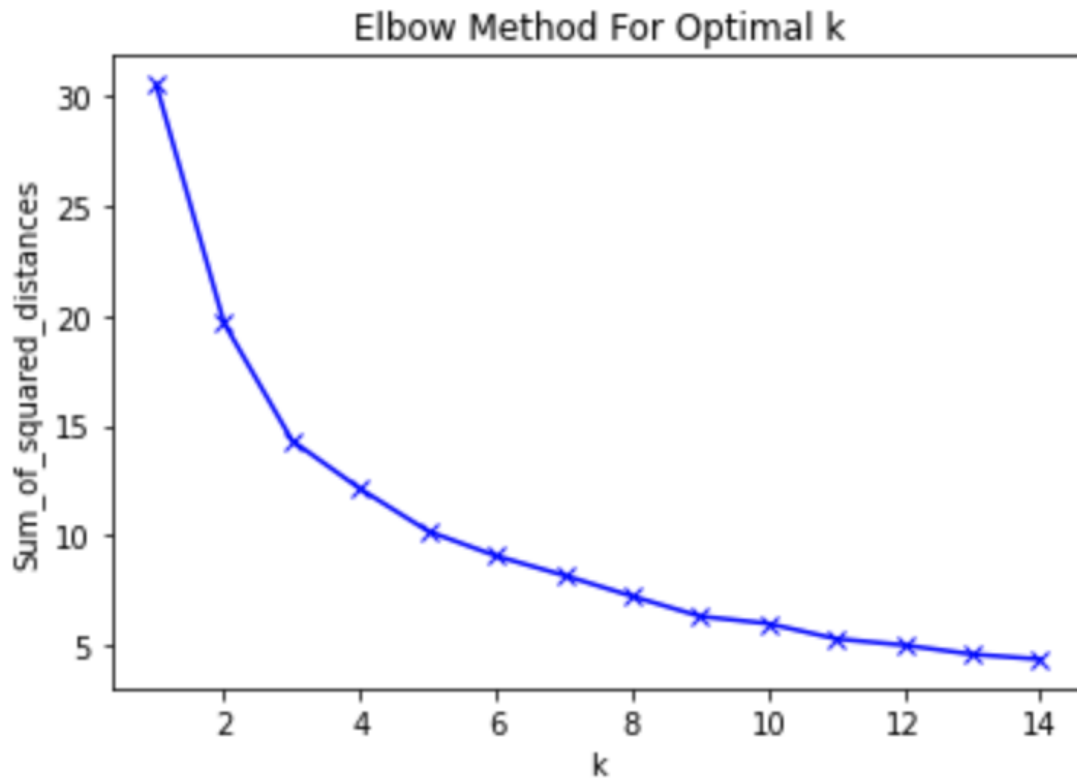
	Postcode district	Coverage	Local Authority Areas	Latitude	Longitude	Arts & Entertainment	College & University	Event	Food	Nightlife Spot	Outdoors & Recreation	Professional & Other Places
0	WC1A	New Oxford Street	Camden	51.517165	-0.126810	124	83	2	247	231	97	113
1	WC1B	Bloomsbury, British Museum, Southampton Row	Camden	51.519140	-0.127759	107	95	2	246	245	99	107
2	WC1E	University College London, SOAS	Camden	51.522520	-0.133033	84	75	1	196	159	80	92
3	WC1H	St Pancras, UCL Institute of Education	Camden	51.524755	-0.131389	74	81	0	77	58	85	86
4	WC1N	Russell Square, Great Ormond Street	Camden	51.524460	-0.123254	75	88	0	189	134	76	88
5	WC1R	Gray's Inn	Camden	51.519375	-0.117610	83	75	2	235	206	84	92
6	WC1V	High Holborn	Camden, City of London	51.518660	-0.112102	61	74	1	220	193	74	102

From the boxplot, we could see the basic distribution of venues in different categories. There are various outliers in Professional & Other Places and Residence category.

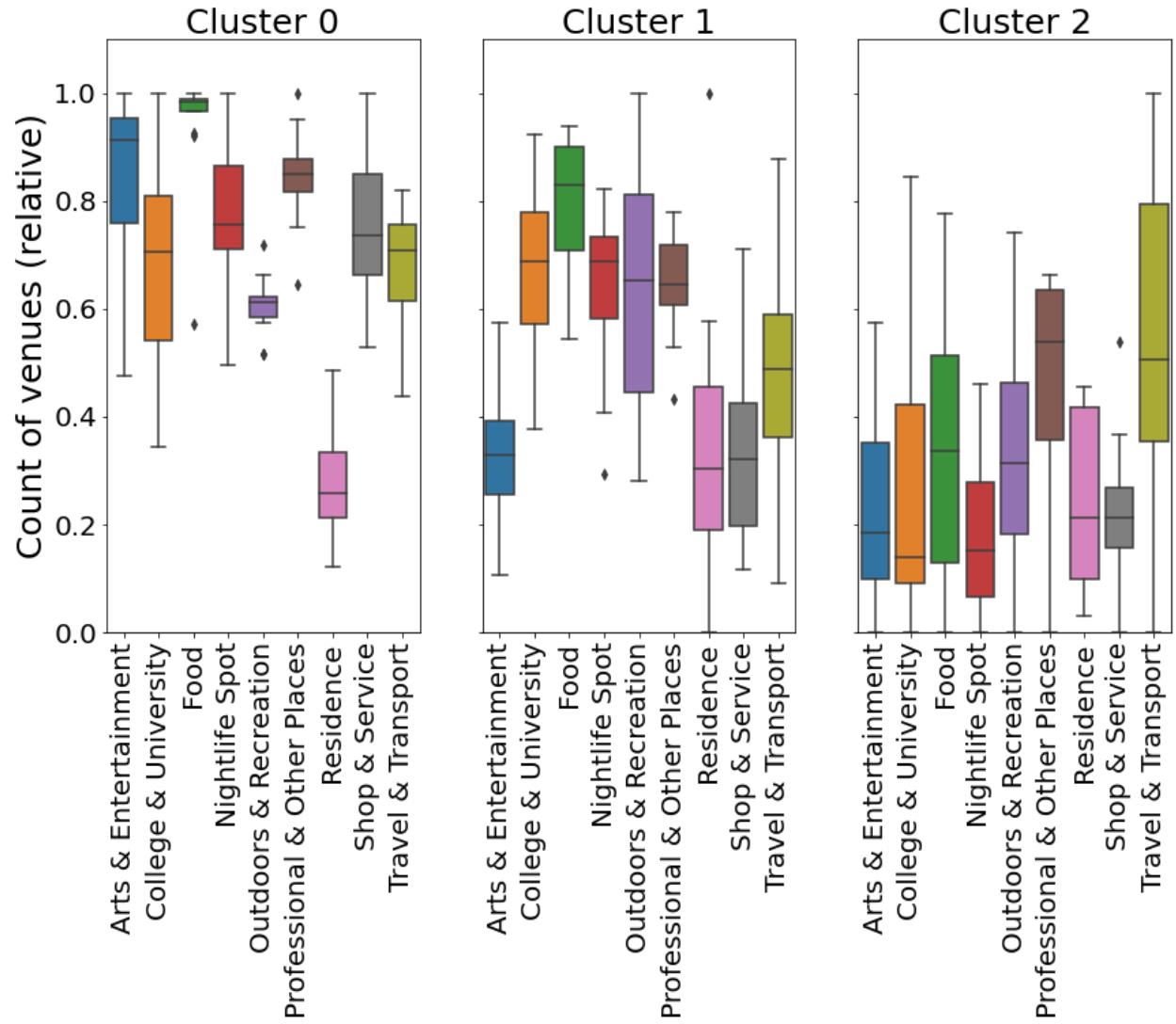


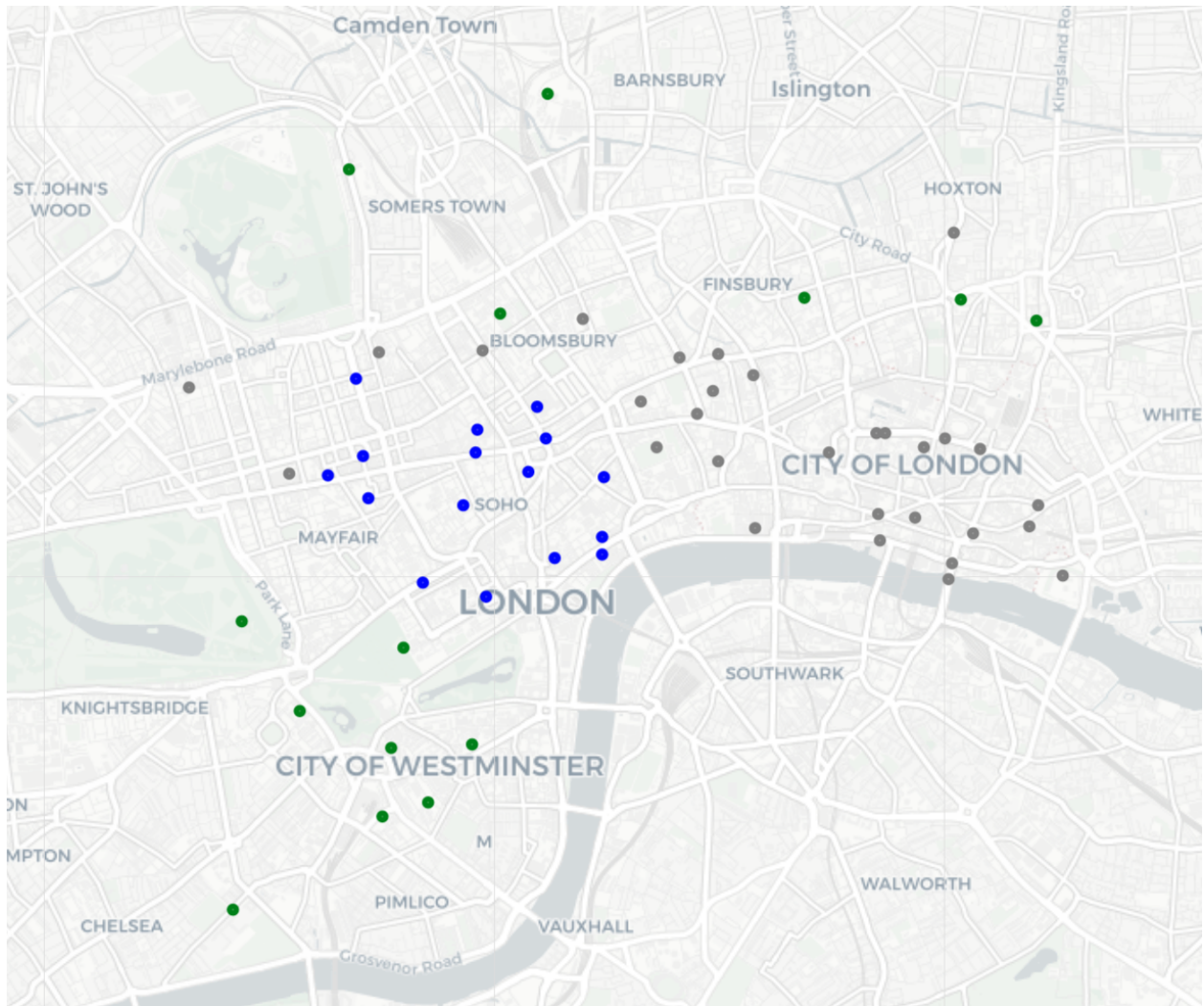
Performing K-means clustering algorithm to segment neighborhoods

First, let's determine the optimal number of clusters



From the elbow graph, we see that 3 clusters are a perfect fit for this study. Following box plots give us a visual representation of clusters.





Conclusion:

Combine the Box Plot and the map, we could see that Cluster 0 (Blue points) apparently has more fun and entertainment places that are full of restaurants, cute shops and theatres. Cluster 1 (Grey points) has less shops but more companies. Cluster 3 (green points) has less fun stuff but acting more as the transportation center in London.