

THE BATTLE OF THE NEIGHBORHOODS

TO SELECT THE BEST PLACE FOR STUDENTS IN LONDON



I. PROBLEM

- London is a **popular destination for higher education** where diverse students from all around the world gather to study.
- According to the data published by the Higher Education Statistical Agency (HESA), in the academic year 2016-2017 London welcomed **112,200 international students** to its higher education institutions, which make up **24 percent of students at higher education institutions**.
- One of the big concerns for international students when moving to a new city would be finding an accommodation
- This project aims to explore different neighborhoods of London and **find the best area to build a new student hall for international students** to solve this persistent problem and to find a new business opportunity.
- A lot of factors come into play when finding the best accommodation, including location and rent. In this project, however, the study focuses only on **the safety and the general atmosphere of the neighborhood**.



2. DATA

(I) London Recorded Crime

- Description: London crime records classified by boroughs and crime type in the last 24 months
- Size: 1584 * 27
- Source: London Datastore, “London Recorded Crime : Geographic Breakdown”
- URL: https://data.london.gov.uk/dataset/recording_crime_summary



London Crime Record

	MajorText	MinorText	LookUp_BoroughName	201707	201708	201709	201710	201711	201712	201801	...	201809	201810	201811
0	Arson and Criminal Damage	Arson	Barking and Dagenham	2	5	8	7	7	4	2	...	3	8	5
1	Arson and Criminal Damage	Criminal Damage	Barking and Dagenham	143	169	134	132	108	119	135	...	107	131	105
2	Burglary	Burglary - Business and Community	Barking and Dagenham	42	30	25	23	27	21	38	...	33	32	39
3	Burglary	Burglary - Residential	Barking and Dagenham	95	83	81	122	88	124	143	...	99	94	106
4	Drug Offences	Drug Trafficking	Barking and Dagenham	7	1	6	7	5	6	4	...	9	6	7

	LookUp_BoroughName	MonthlyAverage
0	Barking and Dagenham	1574.916667
1	Barnet	2395.375000
2	Bexley	1329.708333
3	Brent	2553.958333
4	Bromley	1938.791667



2. DATA

(2) List of London Boroughs

- Description: Information on 32 boroughs of London. ‘Population’ and ‘Coordinates’ will be used for the project.
- Size: 32 * 10
- Source: Wikipedia, “List of London Boroughs”
- URL: https://en.wikipedia.org/wiki/List_of_London_boroughs



LONDON BOROUGHS

Borough	Inner	Status	Local authority	Political control	Headquarters	Area (sq mi)	Population (2013 est) ^[1]	Co-ordinates	Nr. in map
Barking and Dagenham [note 1]			Barking and Dagenham London Borough Council	Labour	Town Hall, 1 Town Square	13.93	194,352	51.5607°N 0.1557°E	25
Barnet			Barnet London Borough Council	Conservative	North London Business Park, Oakleigh Road South	33.49	369,088	51.6252°N 0.1517°W	31
Bexley			Bexley London Borough Council	Conservative	Civic Offices, 2 Watling Street	23.38	236,687	51.4549°N 0.1505°E	23
Brent			Brent London Borough Council	Labour	Brent Civic Centre, Engineers Way	16.70	317,264	51.5588°N 0.2817°W	12
Bromley			Bromley London Borough Council	Conservative	Civic Centre, Stockwell Close	57.97	317,899	51.4039°N 0.0198°E	20
Camden	✓		Camden London Borough Council	Labour	Camden Town Hall, Judd Street	8.40	229,719	51.5290°N 0.1255°W	11

2. DATA

(3) Foursquare API

- Description: List of top 50 popular places in each borough. However, some suburban areas may have less than 50 places registered on Foursquare.
- Size: 1128 * 7
- Source: Foursquare API
- URL: <https://api.foursquare.com>



FOURSQUARE API DATA

	BoroughName	Borough Latitude	Borough Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Barking and Dagenham	51.5607	0.1557	Central Park	51.559560	0.161981	Park
1	Barking and Dagenham	51.5607	0.1557	Beacontree Heath Leisure Centre	51.560997	0.148932	Gym / Fitness Center
2	Barking and Dagenham	51.5607	0.1557	Crowlands Heath Golf Course	51.562457	0.155818	Golf Course
3	Barking and Dagenham	51.5607	0.1557	Robert Clack Leisure Centre	51.560808	0.152704	Martial Arts Dojo
4	Barking and Dagenham	51.5607	0.1557	Morrisons Beacontree Heath	51.559774	0.148752	Supermarket



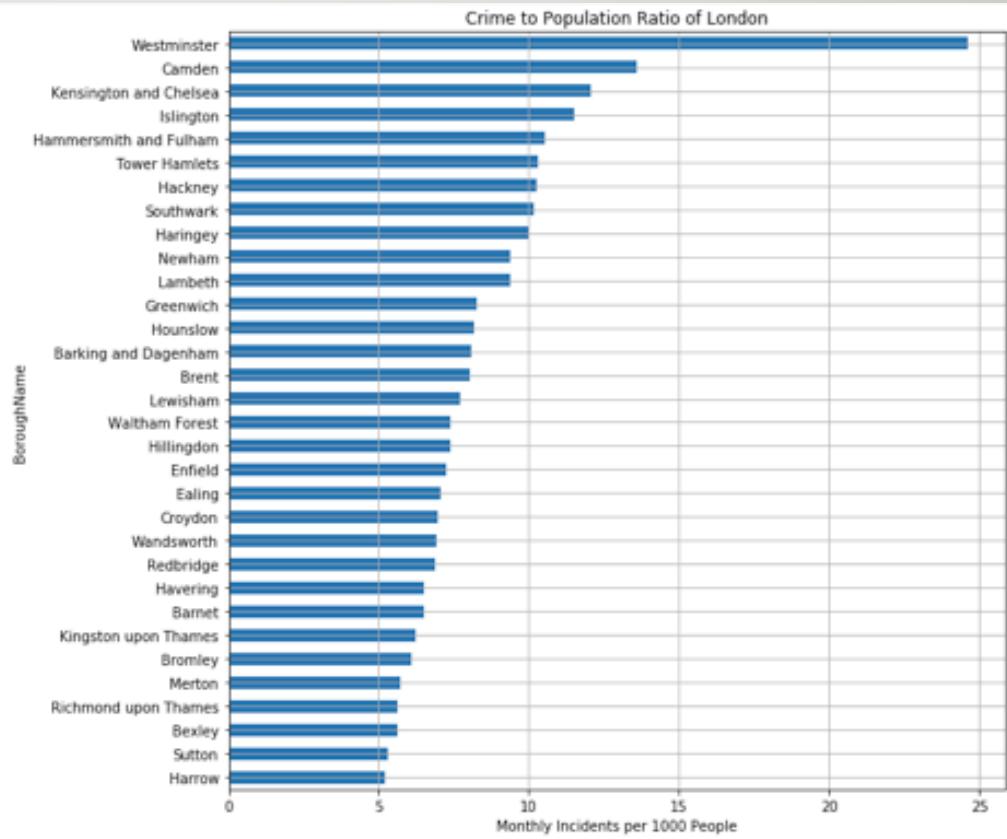
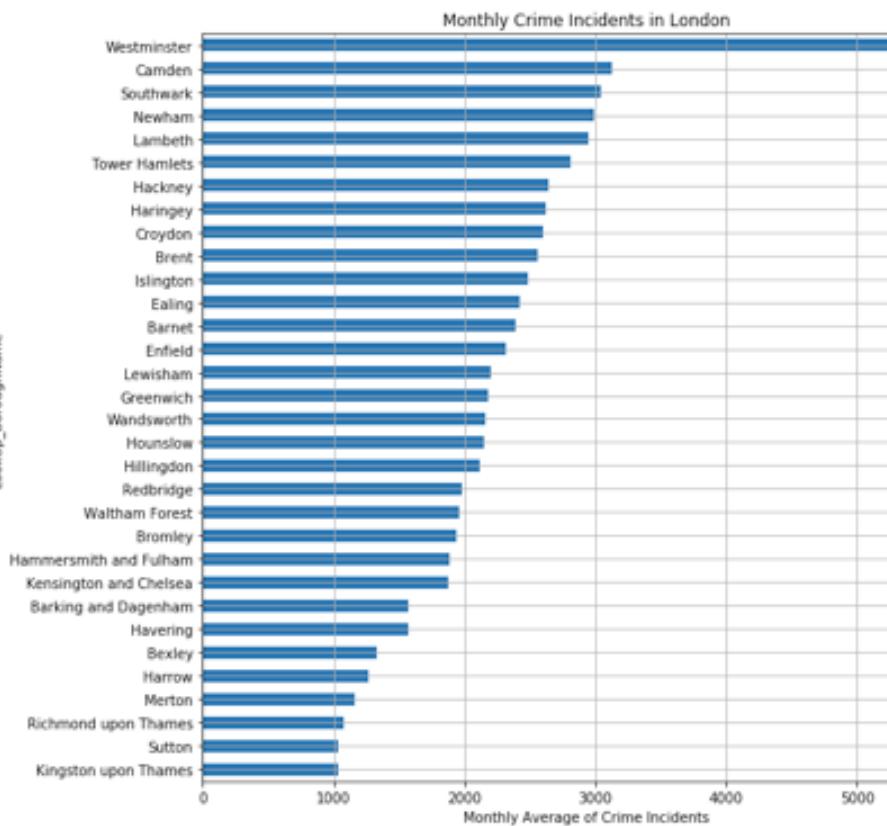
3. METHODOLOGY

(I) Exploratory Analysis: Which borough is more dangerous?

- To observe the datasets better, each datasets were wrangled and cleansed into useful forms.
- Combining the London Recorded Crime data and the List of London Boroughs data, the number of average monthly crime records per 1000 people was calculated.
- When we visualize the boroughs by total number of crimes and by the ratio of crime to population, it is noticeable that the top two dangerous boroughs, **Westminster** and **Camden**, still remain the top two dangerous places



CRIME RATE



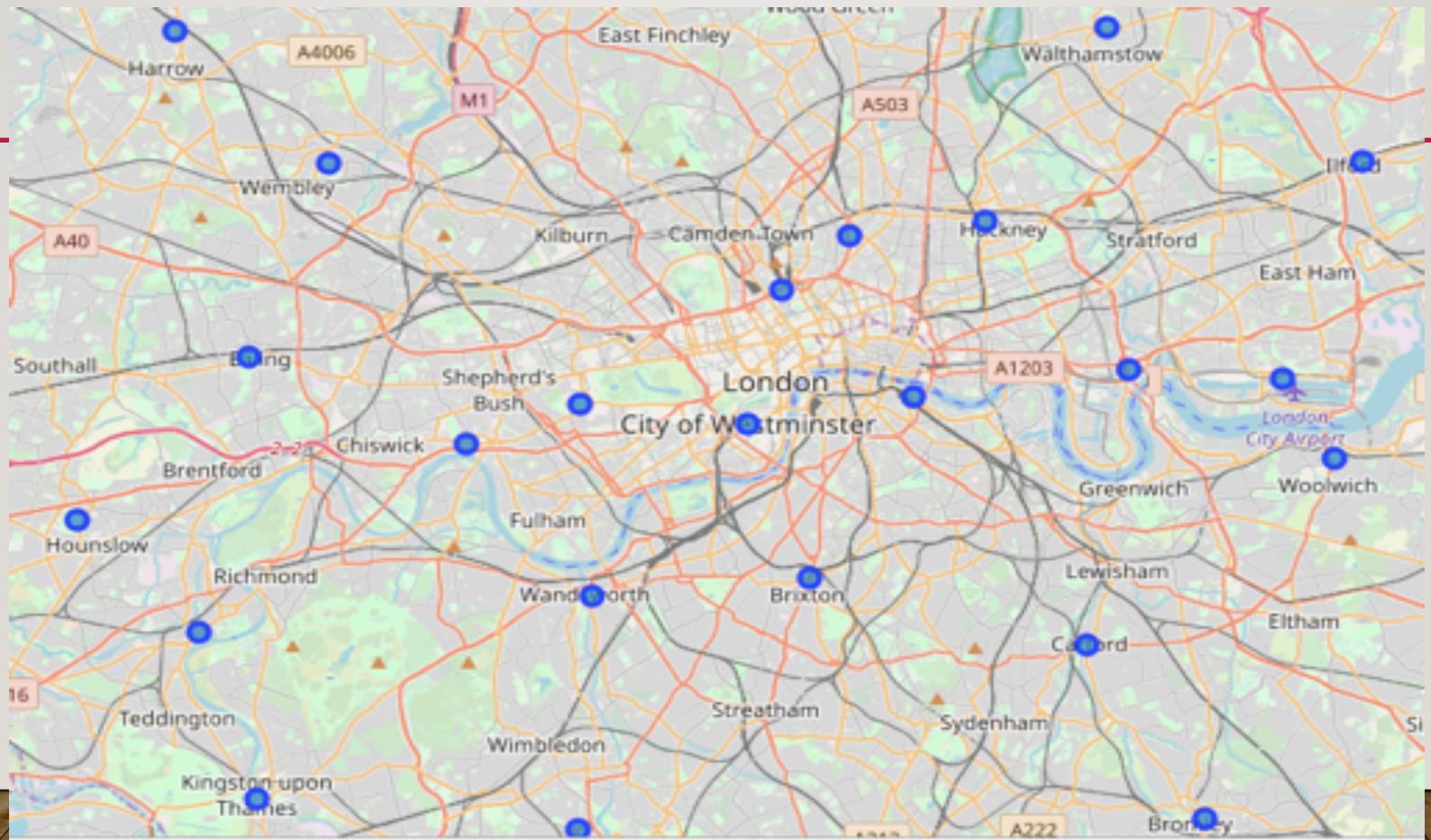
3. METHODOLOGY

(2) Exploratory Analysis: Map of London

- Then with coordinates of each borough from List of London Boroughs data, we could visualize each borough on the map to get familiar with the Greater London area.



GREATER LONDON AREA



3. METHODOLOGY

(3) Cluster Analysis: Popular venues

- From the top 50 venue data we have acquired from Foursquare, we used one hot encoding method to find out what venue categories are most popular.
- Venues from the same boroughs were grouped by borough names and popular categories were discovered by frequency.



COMMON VENUES

	BoroughName	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	Barking and Dagenham	Martial Arts Dojo	Pool	Bus Station	Supermarket	Park	Gym / Fitness Center	Golf Course	Yoga Studio	Flea Market	Fish Market
1	Barnet	Café	Bus Stop	Yoga Studio	Farmers Market	French Restaurant	Food Court	Food	Flea Market	Fish Market	Fish & Chips Shop
2	Bexley	Coffee Shop	Pub	Italian Restaurant	Supermarket	Fast Food Restaurant	Clothing Store	Hotel	Grocery Store	Furniture / Home Store	Multiplex
3	Brent	Coffee Shop	Hotel	Clothing Store	Sporting Goods Shop	Bar	Grocery Store	Sandwich Place	Italian Restaurant	American Restaurant	Café
4	Bromley	Coffee Shop	Clothing Store	Pizza Place	Burger Joint	Bar	Gym / Fitness Center	Noodle House	Stationery Store	Burrito Place	Café



3. METHODOLOGY

(4) Cluster Analysis: K-Means Clustering

- Based on the common venue categories, boroughs were grouped into five clusters as displayed on the map by colors.
- Each clusters were then given a name based on the characteristics they display.



3. METHODOLOGY

- Cluster 0 – TRAVELER AREA

MonthlyAverage	CrimeToPop	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
3	2553.958333	8.049947	0	Coffee Shop	Hotel	Clothing Store	Sporting Goods Shop	Bar	Grocery Store	Sandwich Place
5	3127.916667	13.616273	0	Hotel	Coffee Shop	Café	Train Station	Hotel Bar	Pizza Place	Breakfast Spot
13	1266.500000	5.203968	0	Coffee Shop	Indie Movie Theater	Indian Restaurant	Convenience Store	Platform	Supermarket	Exhibit
23	2989.250000	9.393452	0	Hotel	Airport	Airport Lounge	Airport Service	Chinese Restaurant	Pharmacy	Sandwich Place
28	2810.458333	10.298869	0	Italian Restaurant	Hotel	Coffee Shop	Café	Gym / Fitness Center	Grocery Store	Outdoor Sculpture
31	5584.916667	24.620402	0	Hotel	Sandwich Place	Coffee Shop	Theater	Sushi Restaurant	Hotel Bar	Juice Bar
										Fast Food Restaurant

3. METHODOLOGY

- Cluster 1 – LIVELY AREA

MonthlyAverage	CrimeToPop	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
16	2147.958333	8.185598	1	Bed & Breakfast	Pizza Place	Café	Park	Yoga Studio	Falafel Restaurant	Food Court	Food Market



3. METHODOLOGY

- Cluster 2 – QUIET AREA

MonthlyAverage	CrimeToPop	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	
1	2395.375	6.489983	2	Café	Bus Stop	Yoga Studio	Farmers Market	French Restaurant	Food Court	Food	Flea Market	Fish Market	Fish & Chips Shop



3. METHODOLOGY

- Cluster 3 – BUSY AREA

	MonthlyAverage	CrimeToPop	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
2	1329.708333	5.618003	3	Coffee Shop	Pub	Italian Restaurant	Supermarket	Fast Food Restaurant	Clothing Store	Hotel	Grocery Store
4	1938.791667	6.098766	3	Coffee Shop	Clothing Store	Pizza Place	Burger Joint	Bar	Gym / Fitness Center	Noodle House	Stationery Store
6	2604.333333	6.986772	3	Coffee Shop	Pub	Asian Restaurant	Gym / Fitness Center	Spanish Restaurant	Malay Restaurant	Bookstore	Clothing Store
7	2425.333333	7.081389	3	Coffee Shop	Italian Restaurant	Burger Joint	Park	Bar	Bakery	Pizza Place	Vietnamese Restaurant
8	2322.166667	7.244907	3	Coffee Shop	Sandwich Place	Clothing Store	Pub	Bookstore	Supermarket	Mobile Phone Shop	Café
9	2185.250000	8.277211	3	Coffee Shop	Fast Food Restaurant	Sandwich Place	Pub	Hotel	Supermarket	Grocery Store	Clothing Store

3. METHODOLOGY

- Cluster 4 – HEALTHY AREA

	MonthlyAverage	CrimeToPop	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
0	1574.916667	8.103424	4	Martial Arts Dojo	Pool	Bus Station	Supermarket	Park	Gym / Fitness Center	Golf Course	Yoga Studio	Flea Market	Fish Market

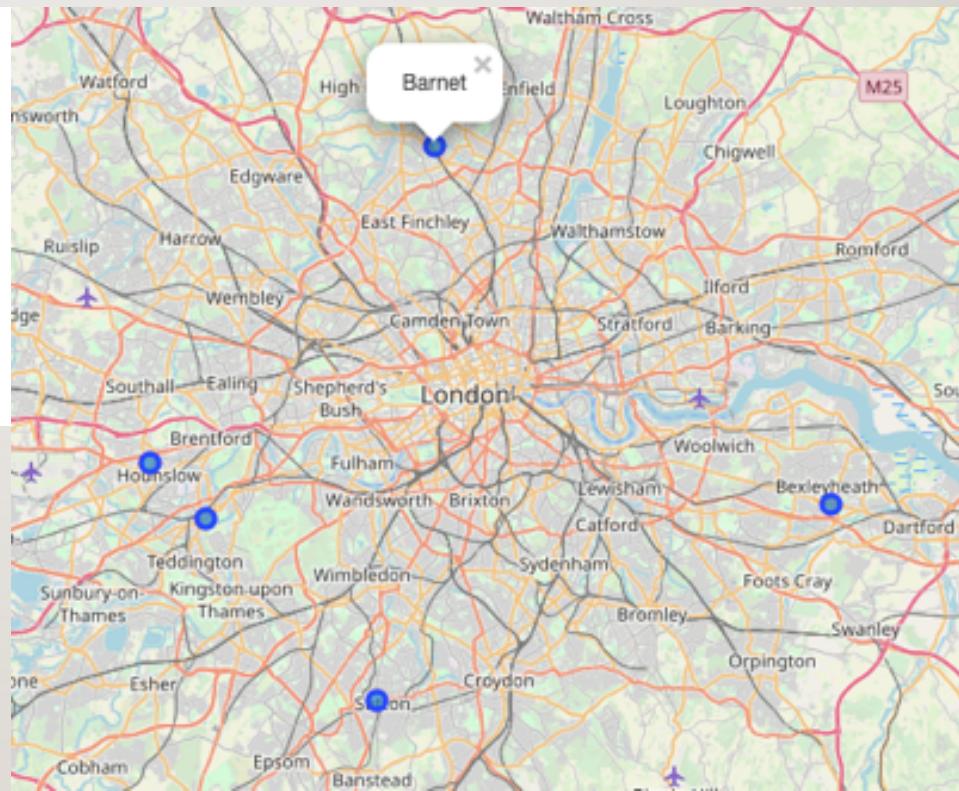
4. RESULTS

- Upon different analysis, we were able to discover the best neighborhoods based on our criteria of safety and atmosphere.
- For safety, we normalized crime to population ratio and reversed the score so that 1 represents the neighborhood with least crime per person.
- For atmosphere, we gave an arbitrary score to each cluster based on personal preference, as preference is not easy to quantify without subjectivity. Highest score was given to Busy area (Cluster 3) which I prefer, and lowest score was given to Traveler area (Cluster 0).
- Then by adding the two scores, we were able to find the best neighborhood scoring highest **Barnet**.



BEST NEIGHBORHOODS

	BoroughName	Safety	Atmosphere	Score
1	Barnet	0.933767	1.0	1.933767
16	Hounslow	0.846438	0.9	1.746438
27	Sutton	0.995057	0.7	1.695057
2	Bexley	0.978676	0.7	1.678676
25	Richmond upon Thames	0.977837	0.7	1.677837



5. CONCLUSION

- However, when we map the top five neighborhoods to live in, it is easily noticeable that they are **all located in far out suburbs**. This is due to many **limitations** this research holds.
- Among numerous factors that determine a good neighborhood, we only took into consideration what type of venues are popular and how many crime incidents are recorded for the sake of simplification.
- To overcome the limitations of this study, **we will need further data** such as distance to city center, housing prices or ratio of the number of stores to population.
- Despite some limitations, this research was still enjoyable in that we were able to explore the neighborhoods in depth.



REFERENCES

Data

- “London Recorded Crime :Geographic Breakdown”, London Datastore
- “List of London Boroughs”, Wikipedia
- Foursquare API

Document

- “The Economic Impact of London’s International Students”, London & Partners (2018)
- Lecture notes from IBM Professional Data Science Specialization, Coursera

