Introduction

1.1 Background

The scenario of this project is to find a relative satest place in London for people who want to live in. As peopel often move from one place to another for the result of study or work, one may need to find a safest place to live beforehand instead of set down randomly. The exploration here give a potential idea where may a good choice for people to live. The information related to crime in a certain area in London can be also used for set up a proper business as safety is a key feature you may look into.

1.2 Interest

People who are moving to live in London for the first time for the reson of study or working may be particularly curious about the safty status of each area. Others who would like to know the general crime rate of each area may also be interested.

2. Data acquisition

2.1 Data source

The data was download from a Kaggle competition (https://www.kaggle.com/jboysen/london-crime (https://www.kaggle.com/jboysen/london-crime)) for crime data in London. It offers the data from 2008 to 2016. To be resonalbe for the result here, we only take the latest data which is year 2016 for analysis.

The data here include the crime types which are:

- Iso_code: code for Lower Super Output Area in Greater London
- borough: Command name for London borough
- major_category: High level categorization for crime
- minor_category: Low level categorization for crime
- value: monthly reported count of categorical crime in given borough
- year: Year of reported counts
- month: Month of reported counts

The list of boroughs was scraped from wiki page (https://en.wikipedia.org/wiki/List_of_London_boroughs (https://en.wikipedia.org/wiki/List_of_London_boroughs)). Following data was offter in the table:

- Borough: London boroughs' names
- Inner: Categorizing as Inner or outer London borough
- Status: Royal, city or other borough
- Local authority: Local authority of that borough
- Political contorl: The political party that control that borough
- Headquarters: Headquarters of the Boroughs
- Area(sq mi): Area of the borough in square miles

- Population: Population number of the borough in year 2013
- Co-ordinate: The latitude and longitude of that borough
- Nr. in map: Number assigned to each borough to represent viually on a map

The other list of neighbohoods of borough Kingston upon Thames was scraped from wiki (https://en.wikipedia.org/wiki/List_of_districts_in_the_Royal_Borough_of_Kingston_upon_Thames (https://en.wikipedia.org/wiki/List_of_districts_in_the_Royal_Borough_of_Kingston_upon_Thames)). The names was used to create a data frmae with the following information:

- Neighbohood: Name of the neighborhood
- Borough: Name of the borough
- Latitude: Latitude of the borough
- Longitude: Longitude of the borough

2.2 Data cleaning

For the final purpose of the data analysis, we need to merge all the data information from different resources into one table. Different resourses have different missing information. The initial London crime data site include data from 2008 to 2016. However, as the time goes, the earlier time may be useful for comparing the crime change by date. For using to check the status now, we only use the data from the cloest year which is 2016. Other data source are scraped by using the wkipedia python package which converted to data frame directly. These data sets are merged into one to visualized the crime rates in each borough and identify the borough with the least crimes recorded during the year 2016.

	Borough	Buglary	Criminal Damage	Drugs	Other Notifiable Offences	Robbery	Theft and Handling	Violence Aginst the Person	Total	Local authority	Political control	Headquarters	Area (sq mi)	Population (2013 est) [1]
0	Barking and Dagenham	1287	1949	919	378	534	5607	6067	16741	Barking and Dagenham London Borough Council	Labour	Town Hall, 1 Town Square	13.93	194352.0
1	Barnet	3402	2183	906	499	464	9731	7499	24684	Barnet London Borough Council	Conservative	North London Business Park, Oakleigh Road South	33.49	369088.0
2	Bexley	1123	1673	646	294	209	4392	4503	12840	Bexley London Borough Council	Conservative	Civic Offices, 2 Watling Street	23.38	236687.0

The second data source which has the boroughs information, which likes the following table:

	Borough	Inner	Status	Local authority	Political control	Headquarters	Area (sq mi)	Population (2013 est)[1]	Co-ordinates	Nr. in map
0	Barking and Dagenham [note 1]	NaN	NaN	Barking and Dagenham London Borough Council	Labour	Town Hall, 1 Town Square	13.93	194352	51°33′39″N 0°09′21″E / 51.5607°N 0.1557°E	25
1	Barnet	NaN	NaN	Barnet London Borough Council	Conservative	North London Business Park, Oakleigh Road South	33.49	369088	51°37′31″N 0°09′06″W / 51.6252°N 0.1517°W	31
2	Bexley	NaN	NaN	Bexley London Borough Council	Conservative	Civic Offices, 2 Watling Street	23.38	236687	51°27′18″N 0°09′02″E / 51.4549°N 0.1505°E	23
3	Brent	NaN	NaN	Brent London Borough Council	Labour	Brent Civic Centre, Engineers Way	16.70	317264	51°33′32″N 0°16′54″W / 51.5588°N 0.2817°W	12
4	Bromley	NaN	NaN	Bromley London Borough Council	Conservative	Civic Centre, Stockwell Close	57.97	317899	51°24′14″N 0°01′11″E / 51.4039°N 0.0198°E	20

Merging all the information ends up with the following dataframe:

	Borough	Buglary	Criminal Damage	Drugs	Other Notifiable Offences	Robbery	Theft and Handling	Violence Aginst the Person	Total	Local authority	Political control	Headquarters	Area (sq mi)	Population (2013 est) [1]
0	Barking and Dagenham	1287	1949	919	378	534	5607	6067	16741	Barking and Dagenham London Borough Council	Labour	Town Hall, 1 Town Square	13.93	194352.0
1	Barnet	3402	2183	906	499	464	9731	7499	24684	Barnet London Borough Council	Conservative	North London Business Park, Oakleigh Road South	33.49	369088.0
2	Bexley	1123	1673	646	294	209	4392	4503	12840	Bexley London Borough Council	Conservative	Civic Offices, 2 Watling Street	23.38	236687.0

3. Methodology

3.1 Exploratory data analysis

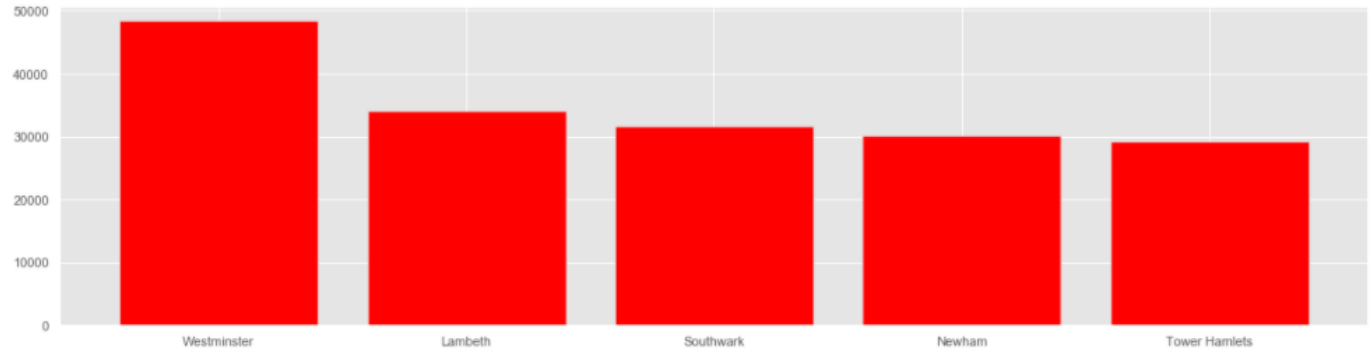
3.1.1 Statistical summary of crimes

Pandas describe function is used to briefing the general statistical information. It returns the mean, standard deviation, minimim, maximum, 1st quartile(25%), 2nd quartile(50%), and 3rd quartile(75%) for each of the major categories of cirme. 'Theft and Handling' is the hightest reported crime and the lowest is 'Other Notifiable offenses'.

	Buglary	Criminal Damage	Drugs	Other Notifiable Offences	Robbery	Theft and Handling	Violence Aginst the Person	Total	Area (sq mi)	Population (2013 est)[1]
count	33.000000	33.000000	33.000000	33.000000	33.000000	33.000000	33.000000	33.000000	33.000000	32.000000
mean	2069.242424	1941.545455	1179.212121	479.060606	682.666667	8913.121212	7041.848485	22306.696970	18.408485	262777.718750
std	737.448644	625.207070	586.406416	223.298698	441.425366	4620.565054	2513.601551	8828.228749	12.645367	57346.611409
min	2.000000	2.000000	10.000000	6.000000	4.000000	129.000000	25.000000	178.000000	1.120000	155594.000000
25%	1531.000000	1650.000000	743.000000	378.000000	377.000000	5919.000000	5936.000000	16903.000000	10.360000	224047.500000
50%	2071.000000	1989.000000	1063.000000	490.000000	599.000000	8925.000000	7409.000000	22730.000000	14.520000	263697.000000
75%	2631.000000	2351.000000	1617.000000	551.000000	936.000000	10789.000000	8832.000000	27174.000000	21.780000	311447.500000
max	3402.000000	3219.000000	2738.000000	1305.000000	1822.000000	27520.000000	10834.000000	48330.000000	57.970000	372752.000000

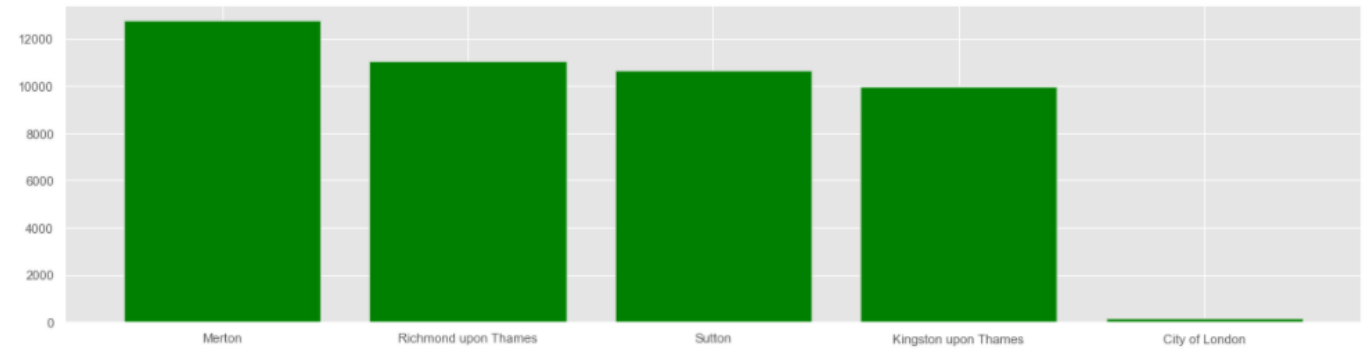
3.1.2 Borough with the highest crime rates

Westminster has the highest crimes recorded followed by Lambeth, Southward, Newham and Tower Hamlets. Westminster has a significantly higher crime rate than the other 4 boroughs.



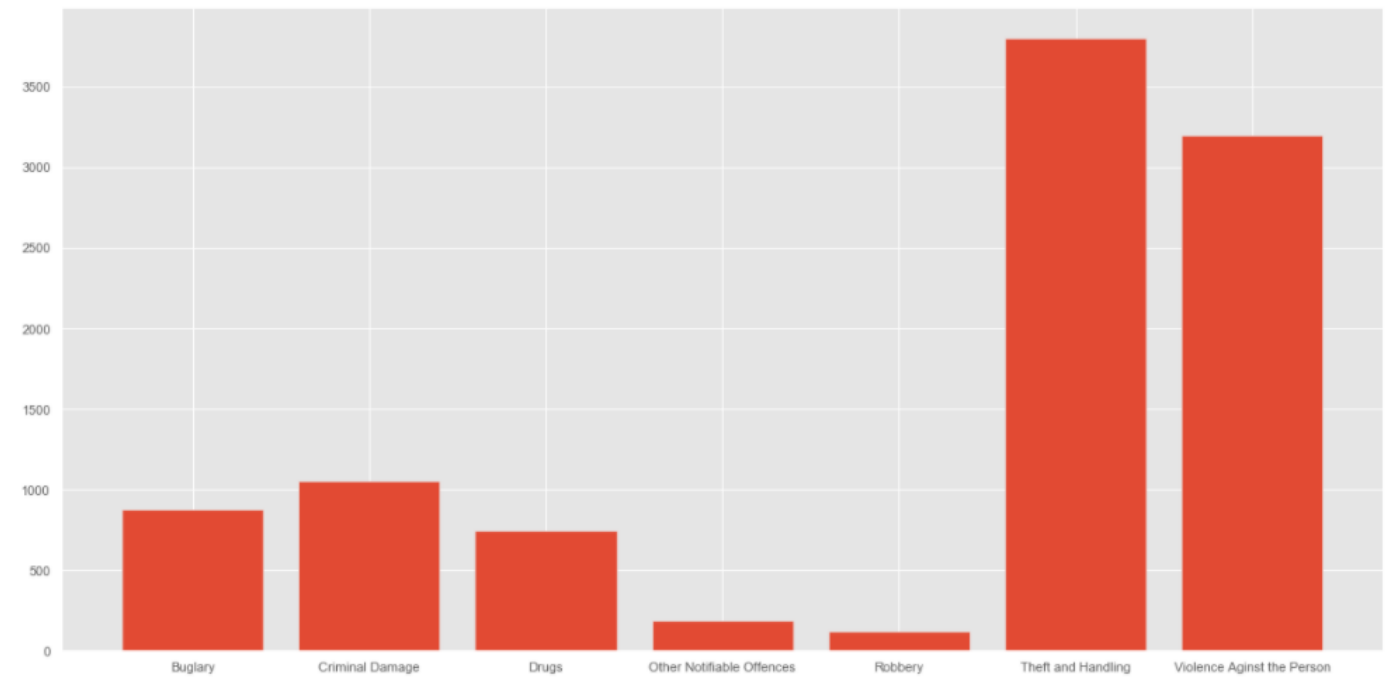
3.1.3 Borough with the lowest crime rates

By comparing the boroughs with the lowest crime rate in 2016, City of London has the lowest record and followd by Kingston upon Thames, Sutton, Richmond upon and Merton.



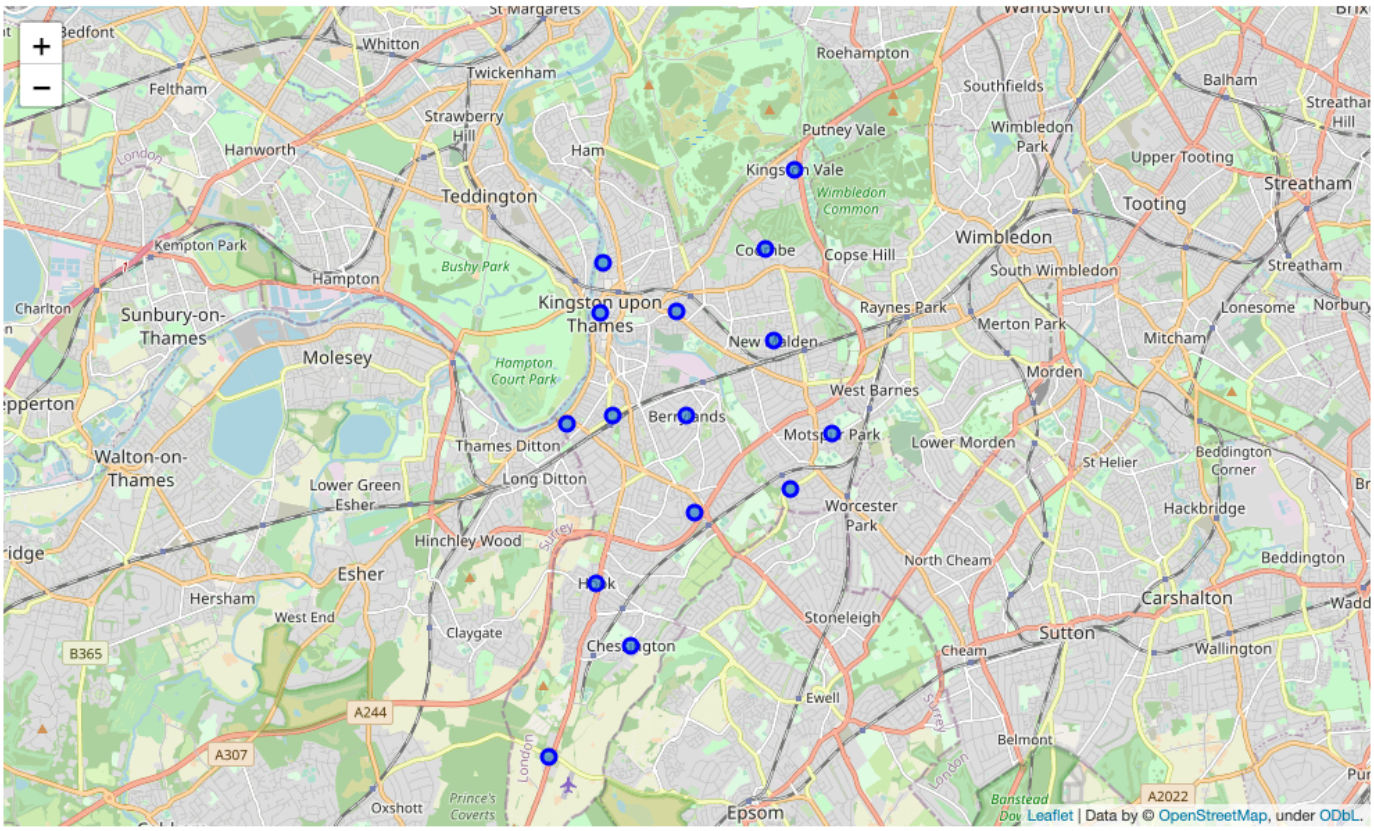
3.1.4 Neighborhoods in Kingston upon Thames

City of London is the principal division of Greater which is not a London borough. Hence we will focus on second to the best borough which is Kingston upon Thames. The different crimes in this borough were show in the following:

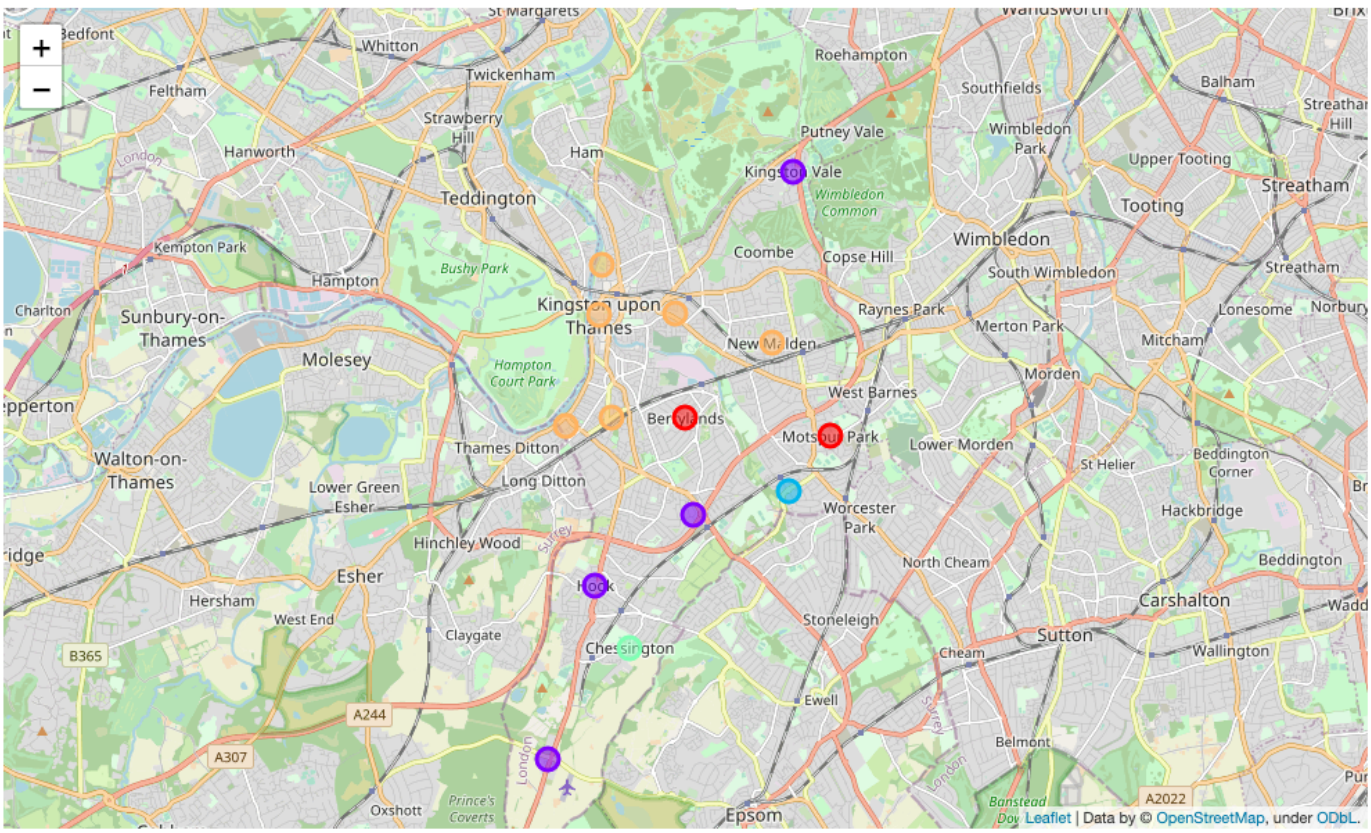


3.1.5 Neighborhoods in Kingston upon Thames

There are 15 neighborhoods in the royal borough of Kingston upon Thames, they are visualised on a map using folium on python.



We have the following picture shows the clustering results.



The majority of the neighborhood falls in the red cluster which is the first. Three neighborhoods have their own cluster(Blue, Purple and Yellow), these clusters two three and five. the green cluster consists of two neighborhoods which is the 4th cluster.

5 Discussion

The aim of this project is to help people who want to relocate to the safest borough in London, expats can chose the neighborhoods to which they want to relocate based on the most common venues in it. For example if a person is looking for a neighborhood with good connectivity and public transportation we can see that Clusters 3 and 4 have Train stations and Bus stops as the most common venues. If a person is looking for a neighborhood with stores and restaurants in a close proximity then the neighborhoods in the first cluster is suitable. For a family I feel that the neighborhoods in Cluster 4 are more suitable dues to the common venues in that cluster, these neighborhoods have common venues such as Parks, Gym/Fitness centers, Bus Stops, Restaurants, Electronics Stores and Soccer fields which is ideal for a family.

6 Conclusion

This project helps a person get a better understanding of the neighborhoods with respect to the most common venues in that neighborhood. It is always helpful to make use of technology to stay one step ahead i.e. finding out more about places before moving into a neighborhood. We have just taken safety as a primary concern to shortlist the borough of London. The future of this project includes taking other factors such as cost of living in the areas into consideration to shortlist the borough based on safety and a predefined budget.