

Battle of Neighborhoods

Capstone Project

Sri Harsha M V - 09 Jun 2020

Introduction

Moving to a new city is probably one of the most glorified tropes of our generation. It's encouraged, if not expected, for people from different ages to head out into the world, to a new city and a new culture, to *find themselves*.

Moving to a new city involves a lot of questions

- How safe is the place ?
- What restaurants do I have around ?
- How easy is it to commute etc

Among these safety is a top concern, if you don't feel safe you won't enjoy staying there.

This project aims to select the safest borough in London based on the total crimes. This project also helps explore the neighbourhood and find out the common venues for each of them.

Data Used

- the data acquired for this project is a combination of data from three sources:
 - The first set of data uses a London crime data which shows the details of all the crimes committed in a borough in London
 - The second source of data is scraped from a Wikipedia page that contains the London boroughs
 - The third data source is the list of neighbourhoods in the Roay Borough of Kingston upon Thames as found on the wikipedia page.

Data Cleansing

- From the London crime data, only the crimes committed during the most recent year are selected. The major categories of crime are pivoted to get the total crimes per the boroughs for each major category
- The second data is scraped from a wikipedia page using the beautiful soup library in python. The data in the table as shown in the webpage is extracted.
- The two data sets are merged on the Borough names to form a new data set. The purpose of this data set is to visualise the crime rates in each borough and identify the borough with the least crimes recorded during the year 2016
- After visualising the crime in each borough we can find the borough with the lowest crime rate. the third data set is created with the names of the neighbourhoods and the name of the borough with the latitude and longitude obtained using google maps API geo coding.
- The new data set is used to generate the 10 most common venues for each neighbourhood using the Foursquare API. Finally using means clustering algorithm to cluster similar neighbourhoods together.

Analysis & Approach

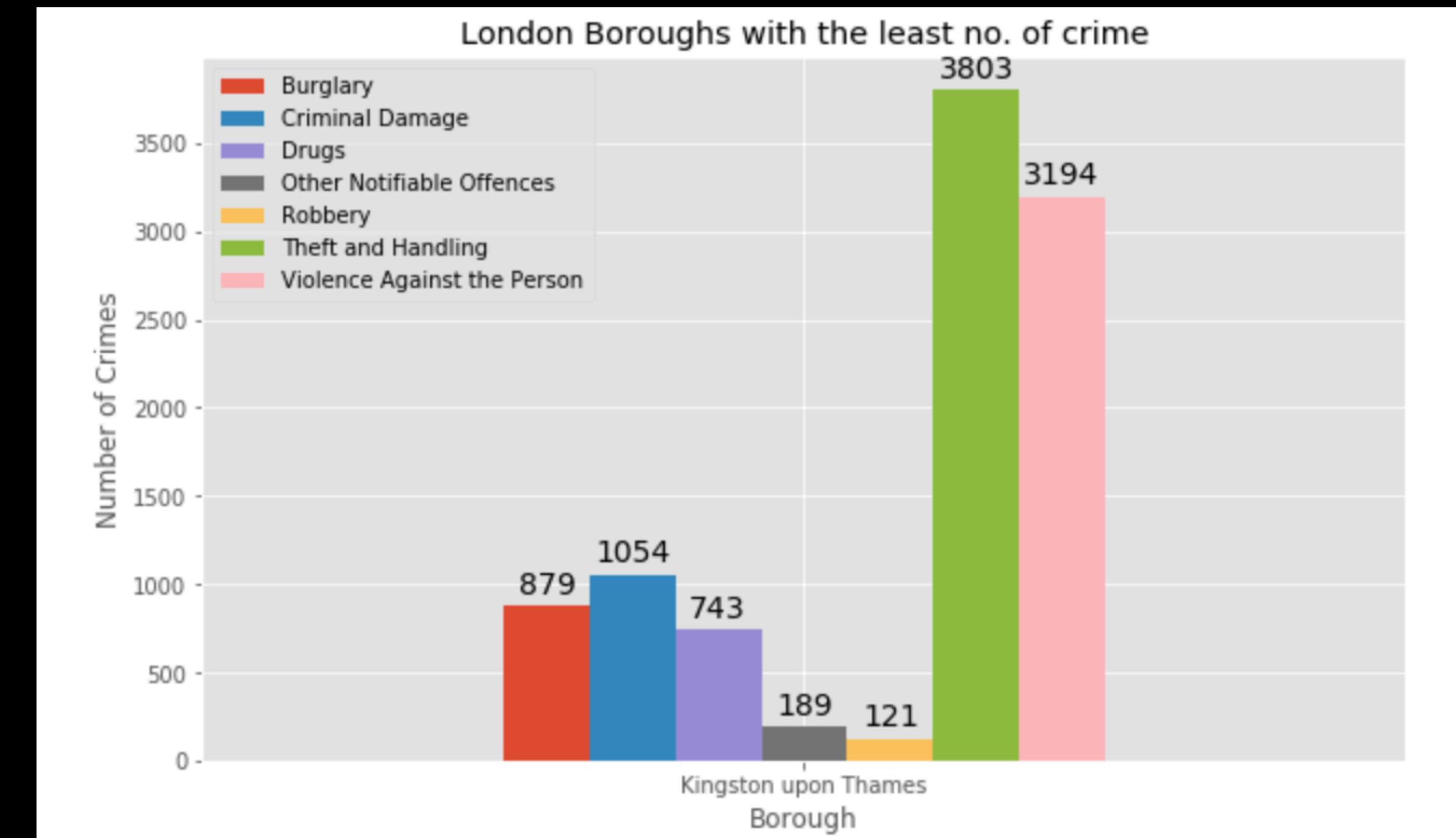
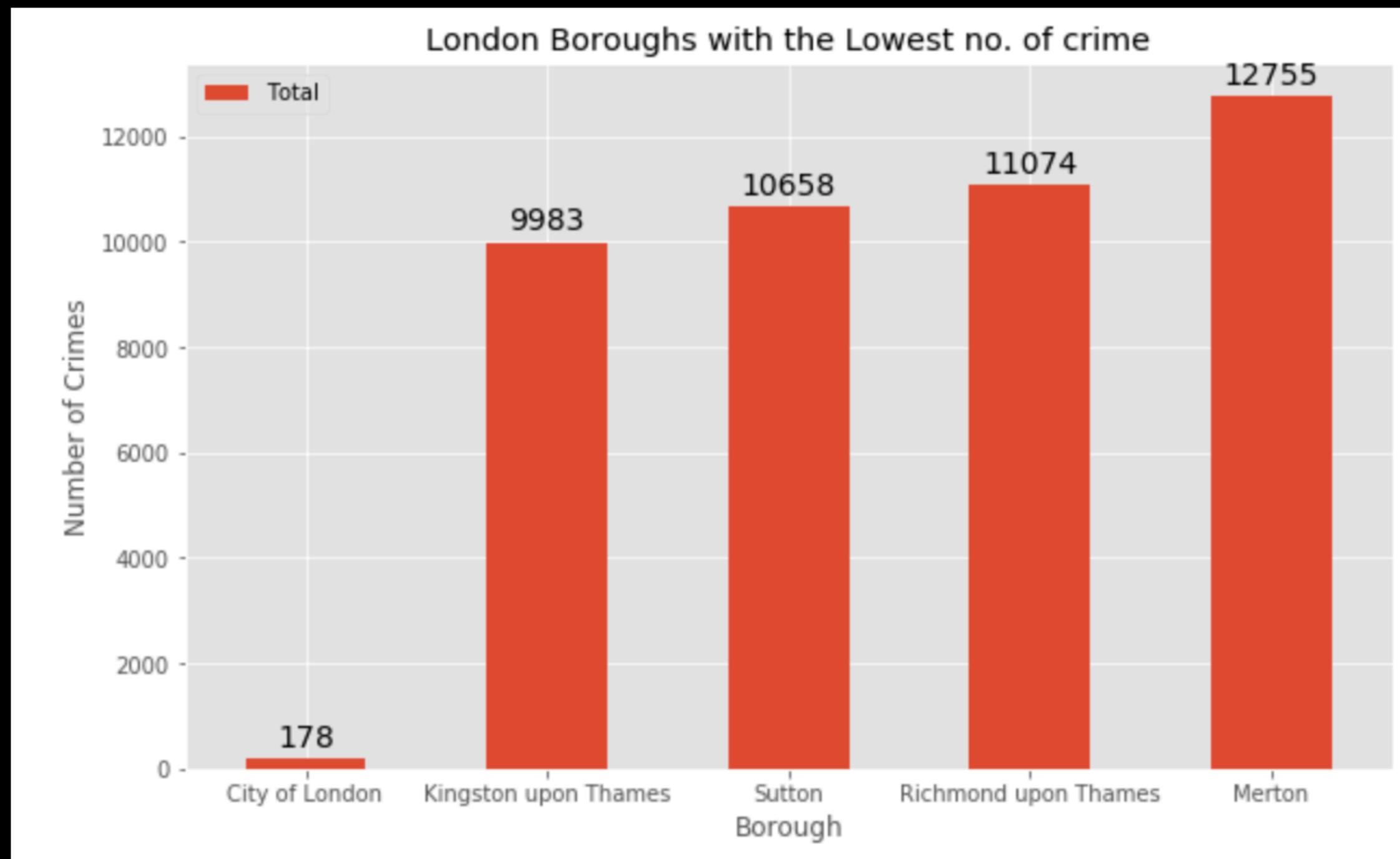
Exploratory Data analysis

Theft and Handling	129159
Violence Against the Person	123050
Criminal Damage	48584
Burglary	43020
Drugs	21782
Robbery	14889
Other Notifiable Offences	11558
Name: Major_Category, dtype: int64	

Major_Category	No_of_Crimes						
	Borough						
	Barking and Dagenham	1287	1949	919	378	534	5607
	Barnet	3402	2183	906	499	464	9731
	Bexley	1123	1673	646	294	209	4392
	Brent	2631	2280	2096	536	919	9026
	Bromley	2214	2202	728	417	369	7584
							6650

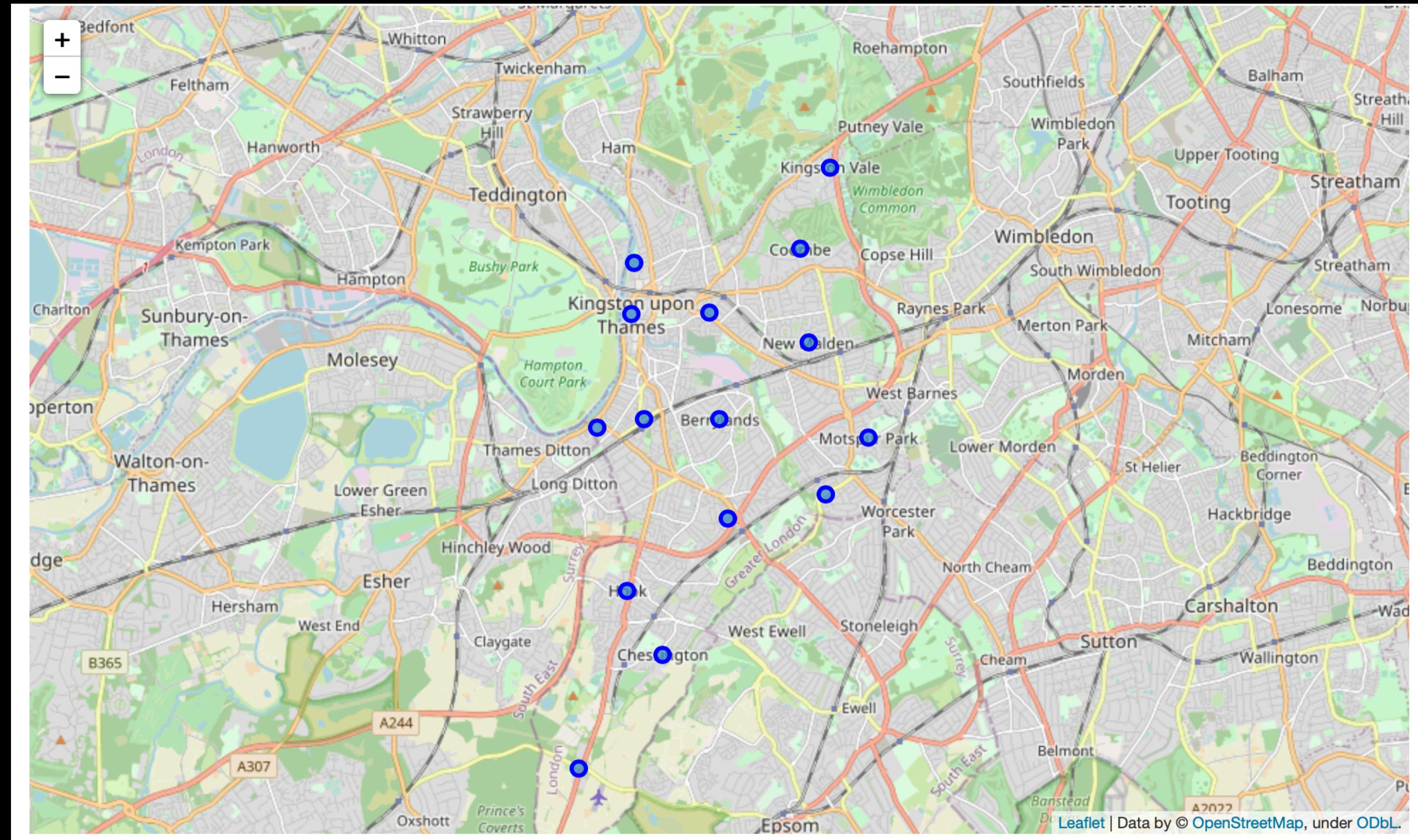
the count for each of the major categories of crime returns the value 33 which is the number of London boroughs. ‘Theft and Handling’ is the highest reported crime during the year 2016 followed by ‘Violence against the person’. The lowest recorded crimes are of ‘Drugs’, ‘Robbery’ and ‘Other Notifiable offences’.

Boroughs with the lowest crimes



- Comparing the 5 boroughs with the lowest crime rate during the year 2016, City of London has the lowest crimes followed by Kingston upon Thames, Sutton, Richmond upon Thames and Merton.
- City of London has a significantly lower crime rate because it is the 33rd principal division of greater London but it is not a London borough.
- Kingston upon Thames is considered as the safest borough in London.

Neighbourhoods in Kingston upon Thames



- There are 15 neighbourhoods in the Kingston upon Thames borough
- These are visualised on a map using Folium in python

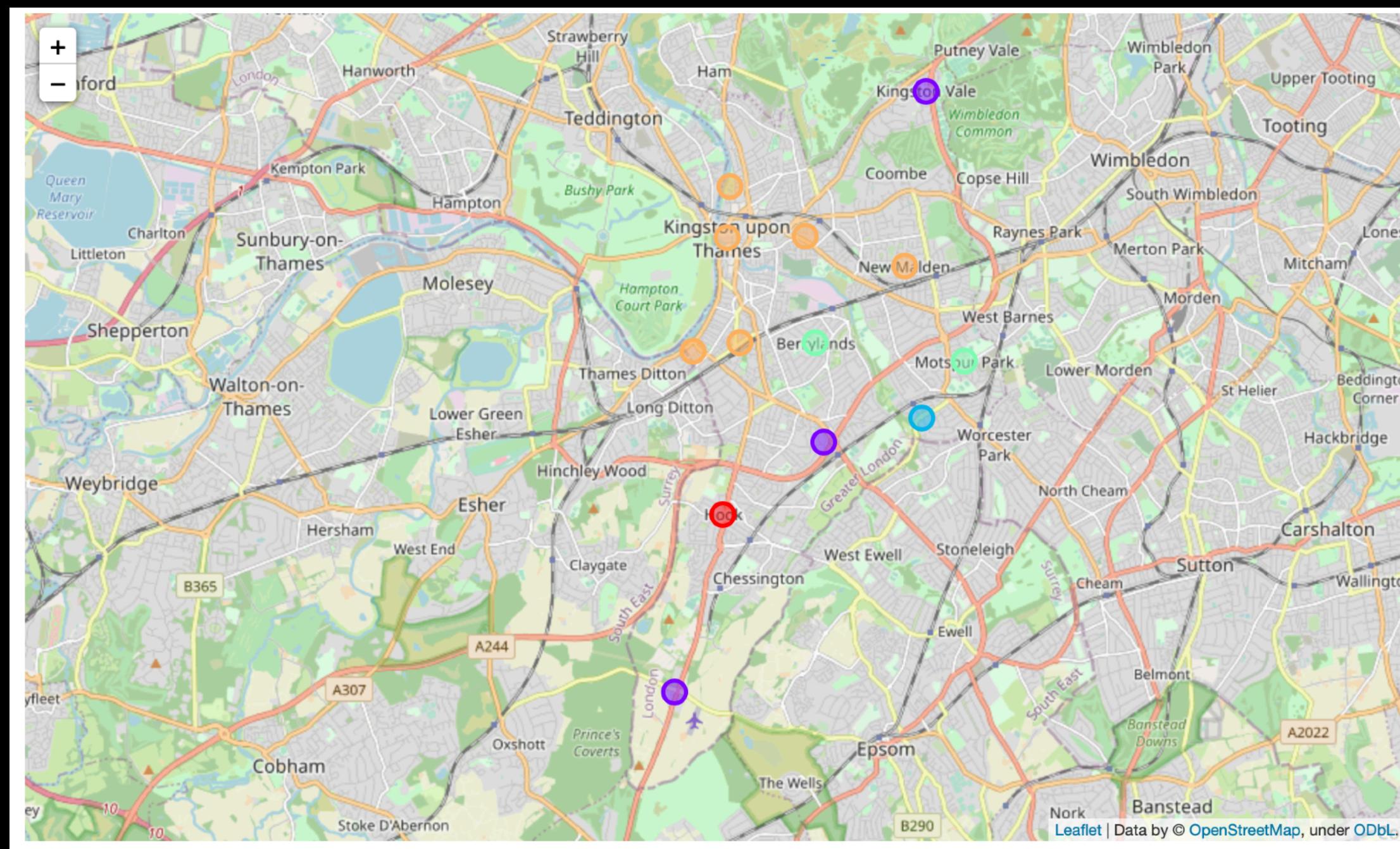
Model

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude		Venue	Venue Latitude	Venue Longitude	Venue Category
0	Berrylands	51.393781	-0.284802	Surbiton Racket & Fitness Club		51.392676	-0.290224	Gym / Fitness Center
1	Berrylands	51.393781	-0.284802		Alexandra Park	51.394230	-0.281206	Park
2	Berrylands	51.393781	-0.284802		K2 Bus Stop	51.392302	-0.281534	Bus Stop
3	Canbury	51.417499	-0.305553		Canbury Gardens	51.417409	-0.305300	Park
4	Canbury	51.417499	-0.305553		The Boater's Inn	51.418546	-0.305915	Pub

- Using the final data set containing the neighbourhoods in Kingston upon Thames alone with their latitude and longitude, we can find all the venues within a 500m radius of each neighbourhood by connecting to the Foursquare API
- One hot encoding is performed on the venues data, which is then grouped by the neighbourhood and the mean of the venues are calculated.
- Subsequently, the 10 common venues are calculated for each of the neighbourhoods.
- To help people find similar neighbourhoods in the safest borough we will be clustering similar neighbourhoods using K-means clustering which is a form of unsupervised machine learning algorithm that clusters data based on predefined cluster size.
- We have used a cluster size of 5 for this project, to split the 15 neighbourhoods into 5 groups.

Observations

- After K-means clustering on the data, we can observe neighbourhoods and their clusters. These are visualised using Folium in python



- Each cluster is color coded for ease of presentation, it can be observed that the orange cluster has majority of the neighbourhoods.

Cluster 1														
Neighborhood	Borough	Latitude	Longitude	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	
4	Hook	Kingston upon Thames	51.367898	-0.307145	0.0	Indian Restaurant	Bakery	Supermarket	Fish & Chips Shop	Food	Department Store	Discount Store	Electronics Store	Farm Ma

Cluster 1

Cluster 5													
Neighborhood	Borough	Latitude	Longitude	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	venue	venue	venue
1	Canbury	Kingston upon Thames	51.417499	-0.305553	4.0	Pub	Fish & Chips Shop	Supermarket	Spa	Gym / Fitness Center	Shop & Service	Plaza	Park
5	Kingston upon Thames	Kingston upon Thames	51.409627	-0.306262	4.0	Café	Coffee Shop	Sushi Restaurant	Burger Joint	Pub	Portuguese Restaurant	French Restaurant	German Restaurant
9	New Malden	Kingston upon Thames	51.405335	-0.263407	4.0	Gastropub	Gym	Sushi Restaurant	Supermarket	Bar	Chinese Restaurant	Korean Restaurant	Indian Restaurant
10	Norbiton	Kingston upon Thames	51.409999	-0.287396	4.0	Indian Restaurant	Italian Restaurant	Food	Pub	Wine Shop	Breakfast Spot	Farmers Market	Hardware Store
12	Seething Wells	Kingston upon Thames	51.392642	-0.314366	4.0	Indian Restaurant	Coffee Shop	Pub	Fast Food Restaurant	Chinese Restaurant	Italian Restaurant	Café	Restaurant
13	Surbiton	Kingston upon Thames	51.393756	-0.303310	4.0	Coffee Shop	Pub	Grocery Store	Italian Restaurant	Pharmacy	Train Station	Hotel	Farmers Market

Cluster 2

Neighborhood	Borough	Latitude	Longitude	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	
					Common									
6	Kingston Vale	Kingston upon Thames	51.431850	-0.258138	1.0	Grocery Store	Sandwich Place	Bar	Soccer Field	Deli / Bodega	Department Store	Discount Store	Electronics Store	Farmers Market
7	Malden Rushett	Kingston upon Thames	51.341052	-0.319076	1.0	Grocery Store	Garden Center	Pub	Restaurant	Farmers Market	Deli / Bodega	Department Store	Discount Store	Electronics Store
14	Tolworth	Kingston upon Thames	51.378876	-0.282860	1.0	Grocery Store	Pharmacy	Sandwich Place	Train Station	Hotel	Discount Store	Coffee Shop	Café	Business Center

Cluster 3

Cluster 3														
	Neighborhood	Borough	Latitude	Longitude	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
11	Old Malden	Kingston upon Thames	51.382484	-0.25909	2.0	Train Station	Pub	Food	Construction & Landscaping	Fast Food Restaurant	Deli / Bodega	Department Store	Discount Store	Elect

Cluster 4

Neighborhood	Borough	Latitude	Longitude	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
Berrylands	Kingston upon Thames	51.393781	-0.284802	3.0	Gym / Fitness Center	Park	Bus Stop	Wine Shop	Fast Food Restaurant	Department Store	Discount Store	Electronics Store	Farmer's Market
Motspur Park	Kingston upon Thames	51.390985	-0.248898	3.0	Gym	Park	Bus Stop	Soccer Field	Wine Shop	Fast Food Restaurant	Department Store	Discount Store	Electronics Store

- Cluster 4 has majority of the neighbourhoods. Each of the neighbourhoods as a specific type of most visited venues.
 - People intending to move into the Boroughs can take this project as a guide to check the neighbourhoods.

Conclusion

- This project helps us get a better understanding of the neighbourhoods with respect to the most common venues in that neighbourhood. It is always helpful to make use of technology to stay one step ahead 1.e. finding out more about places before moving into a neighbourhood.
- We have just taken safety as a primary concern to shortlist the safest borough of London, The future of the project includes taking the other factors such as cost of living into consideration to short list the borough such as filtering areas based on predefined budget.