

# The Battle of Neighborhoods

## 1. Introduction

**Background :** Safety is the top concern when moving to a new area. So, we have to ensure that it's the top priority for us to find a place where we can feel safe.

**Problem :** This project aims to select the safest borough in LONDON based on the total crimes, exploring the neighborhoods of that borough to find the 10 most common venues in each neighborhood and finally cluster the neighborhoods using K-Means Clustering.

**Interest :** Expats who are considering to relocate in London will be interested to identify the safest borough in LONDON and explore its neighborhoods and common venues around each neighborhoods.

## 2. Data Acquisition and Cleaning

- **Data Acquisition:** The data acquired for this project is the combination of data from [three sources](#):
- The first data source for the project uses a [London crime data](#) that shows the crimes per borough in London.
- The second source of data is scraped from a [Wikipedia](#) page that contains the list of London boroughs. This page contains additional information about the boroughs.
- The third data source is the [list of neighborhoods in the Royal Borough of Kingston upon Thames](#) as found on the Wikipedia page.

## ○ Data Cleaning:

The data cleaning process for each of the three sources are done separately.

- From the London crime data , the crimes during the most recent year (2016) are selected . The major categories of crime are pivoted to get the total crimes per the boroughs for each major category.
- The second data source is scraped from a Wikipedia page using [the Beautiful Soup Library](#) in Python. Using this library we can extract the data in a tabular format as shown in the website.
- The two datasets are merged on the borough names to form a new dataset. The purpose of the dataset is to visualize the crime rates of each borough and identify the borough with the least crimes recorded during the year 2016.
- The third dataset is created with the names of the neighborhoods and the names of the borough with the latitude and longitude obtained using [Google Maps API Geo-coding](#).
- The new dataset is used to generate the 10 most common values for each neighborhood using the [Foursquare API](#), finally using the [K-Means Cluster Algorithm](#) to cluster similar neighborhoods together.

### 3. Methodology

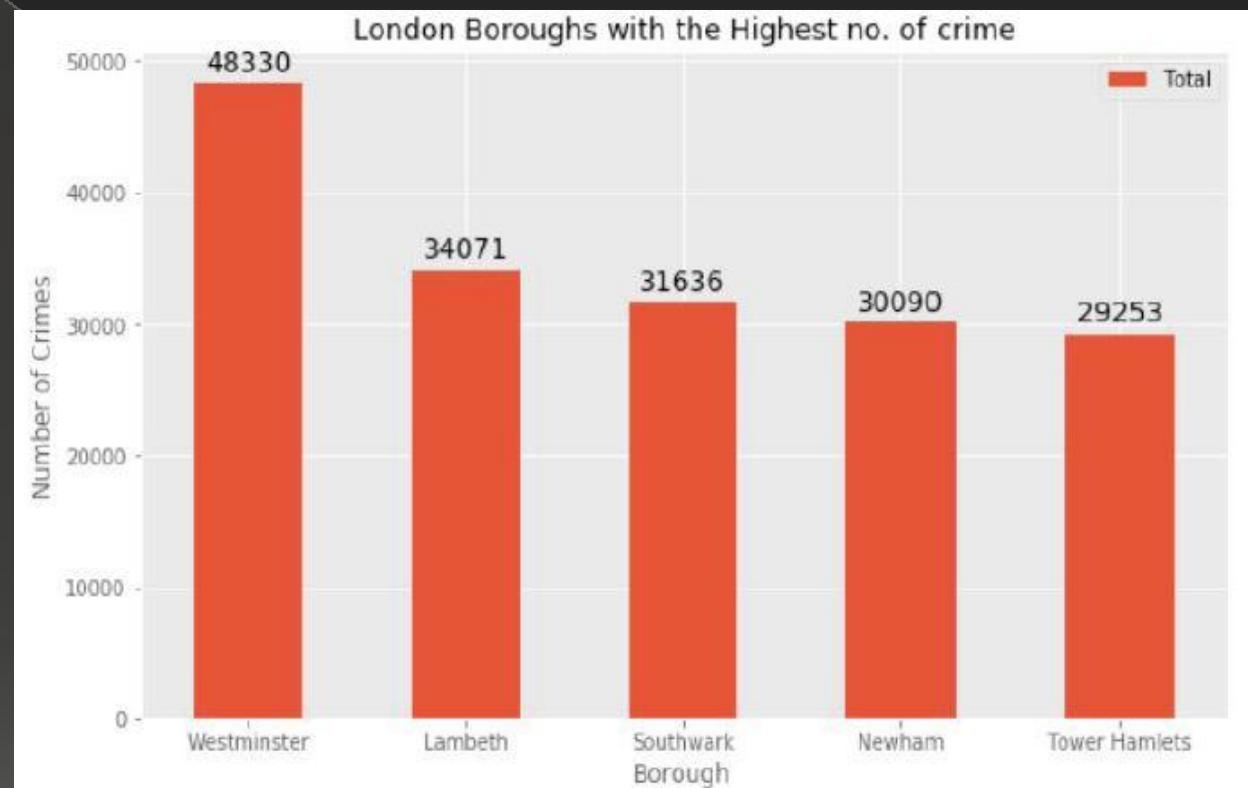
#### ○ 3.1 Exploratory Data Analysis

##### 3.1.1 Statistical summary of crimes

	Burglary	Criminal Damage	Drugs	Other Notifiable Offences	Robbery	Theft and Handling	Violence Against the Person	Total
count	33.000000	33.000000	33.000000	33.000000	33.000000	33.000000	33.000000	33.000000
mean	2069.242424	1941.545455	1179.212121	479.060606	682.666667	8913.121212	7041.848485	22306.696970
std	737.448644	625.207070	586.406416	223.298698	441.425366	4620.565054	2513.601551	8828.228749
min	2.000000	2.000000	10.000000	6.000000	4.000000	129.000000	25.000000	178.000000
25%	1531.000000	1650.000000	743.000000	378.000000	377.000000	5919.000000	5936.000000	16903.000000
50%	2071.000000	1989.000000	1063.000000	490.000000	599.000000	8925.000000	7409.000000	22730.000000
75%	2631.000000	2351.000000	1617.000000	551.000000	936.000000	10789.000000	8632.000000	27174.000000
max	3402.000000	3219.000000	2738.000000	1305.000000	1822.000000	27520.000000	10834.000000	48330.000000

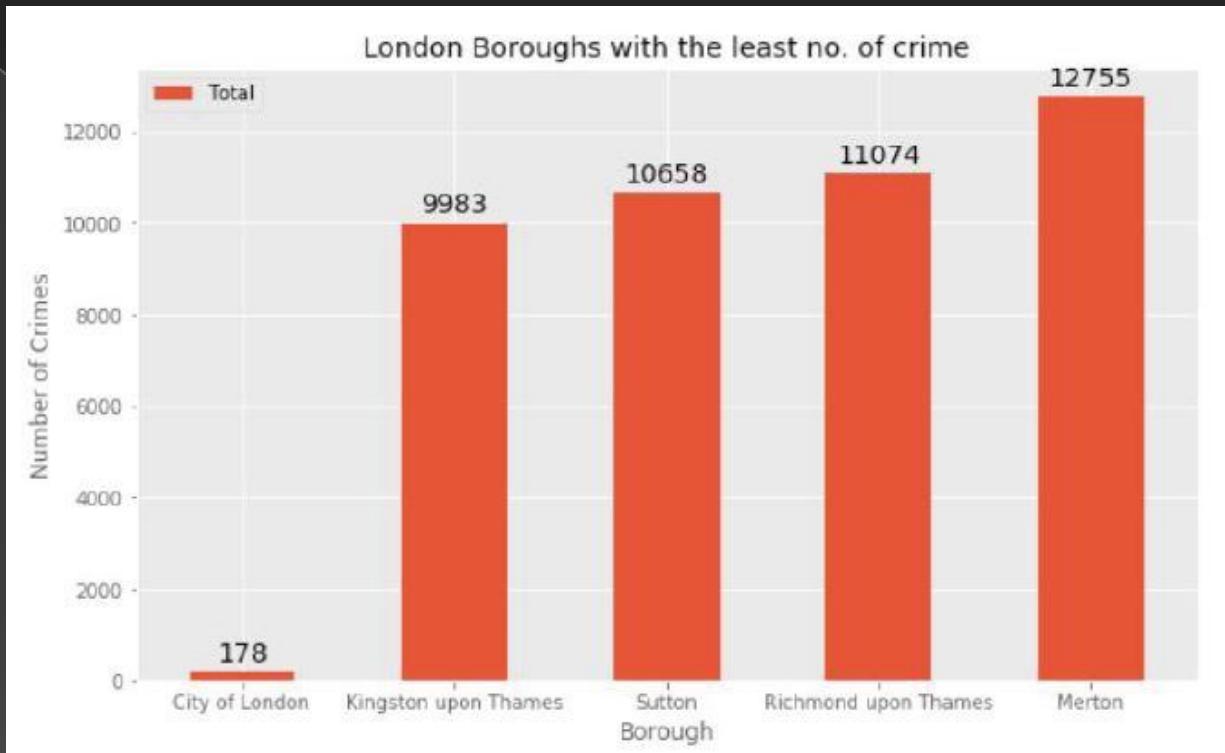
The count for each of the major categories of crime returns the value 33 which is the number of London boroughs. 'Theft and Handling' is the highest reported crime during the year 2016 followed by 'Violence against the person', 'Criminal damage'. The lowest recorded crimes are 'Drugs', 'Robbery' and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

### 3.1.2 Boroughs with the highest crime rates



Comparing five boroughs with the highest crime rate during the year 2016 it is evident that Westminster has the highest crimes recorded followed by Lambeth, Southwark, Newham and Tower Hamlets. Westminster has a significantly higher crime rate than the other 4 boroughs

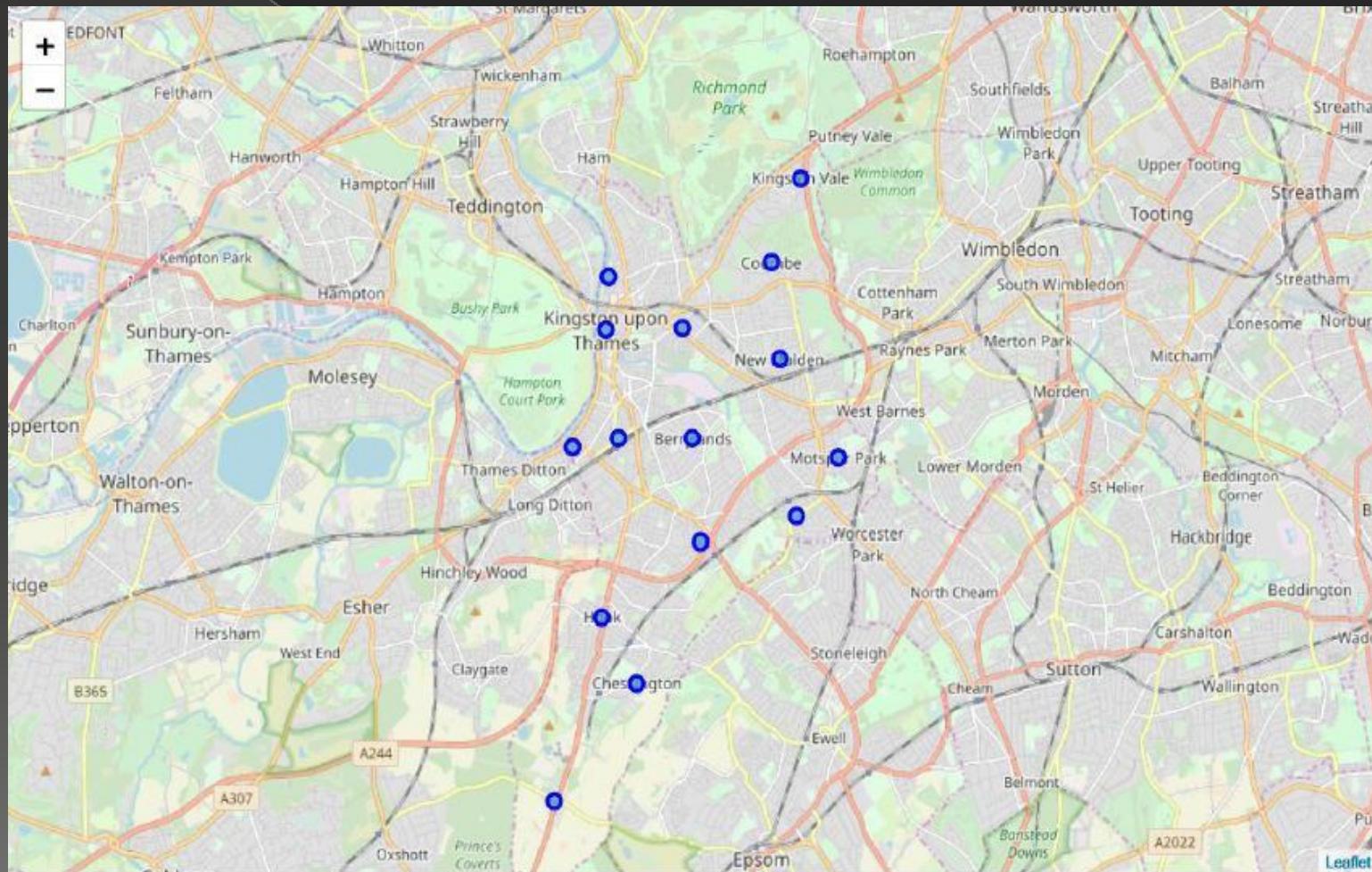
### 3.1.3 Boroughs with the lowest crime rates



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. It has an area of 1.12 square miles and a population of 7000 as of 2013 which suggests that it is a small area ( see fig 3.1.3.1 ). Hence, we will consider the next borough with the lowest crime rate as the safest borough in London which is Kingston upon Thames.

### 3.1.4 Neighborhoods in Kingston upon Thames

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



## 4. Results

- After running the **K-means clustering** we can access each cluster created to see which neighborhoods were assigned to each of the five clusters

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	
0	Berrylands	Kingston upon Thames	51.393781	-0.284802	3	Gym / Fitness Center	Park	Café	Bus Stop	Wine Shop	Fish & Chips Shop	Electronics Store	Farmers Market	Fast Food Restaurant
8	Motspur Park	Kingston upon Thames	51.390985	-0.248898	3	Park	Gym	Restaurant	Soccer Field	Bus Stop	Wine Shop	Fast Food Restaurant	Dry Cleaner	Electronics Store

- The **fourth cluster** has two neighborhoods in it, these neighborhoods have common venues such as Parks, Gym/Fitness centers, Bus Stops, Restaurants, Electronics Stores and Soccer fields etc.

## 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 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 safest 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, such as filtering areas based on a predefined budget.