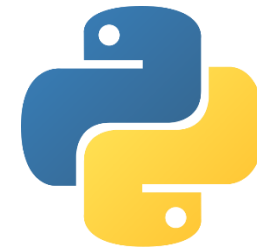


The Battle of Neighbourhoods Presentation



Wesley Gace
July 2020

Introduction

- **Background:** Safety is a top concern when moving to a new area. If you don't feel safe in your own home, you're not going to be able to enjoy living there.
- **Problem:** This project aims to select the safest borough in London, based on the total crimes. This will be done through the use of k-means clustering, which is a type of Machine Learning algorithm.
- **Interest:** People who are considering relocating to London will be interested in identifying the safest borough in the area and common venues around each neighbourhood.

Data Acquisition

The data used in this project is combination of data from three sources:

1. London crimes data that shows crime per borough in London
2. Data scraped from a Wikipedia page that contains a list of London Boroughs
3. The list of neighbourhoods in the Royal Borough of Kingston upon Thames (Wikipedia)

Data Cleaning

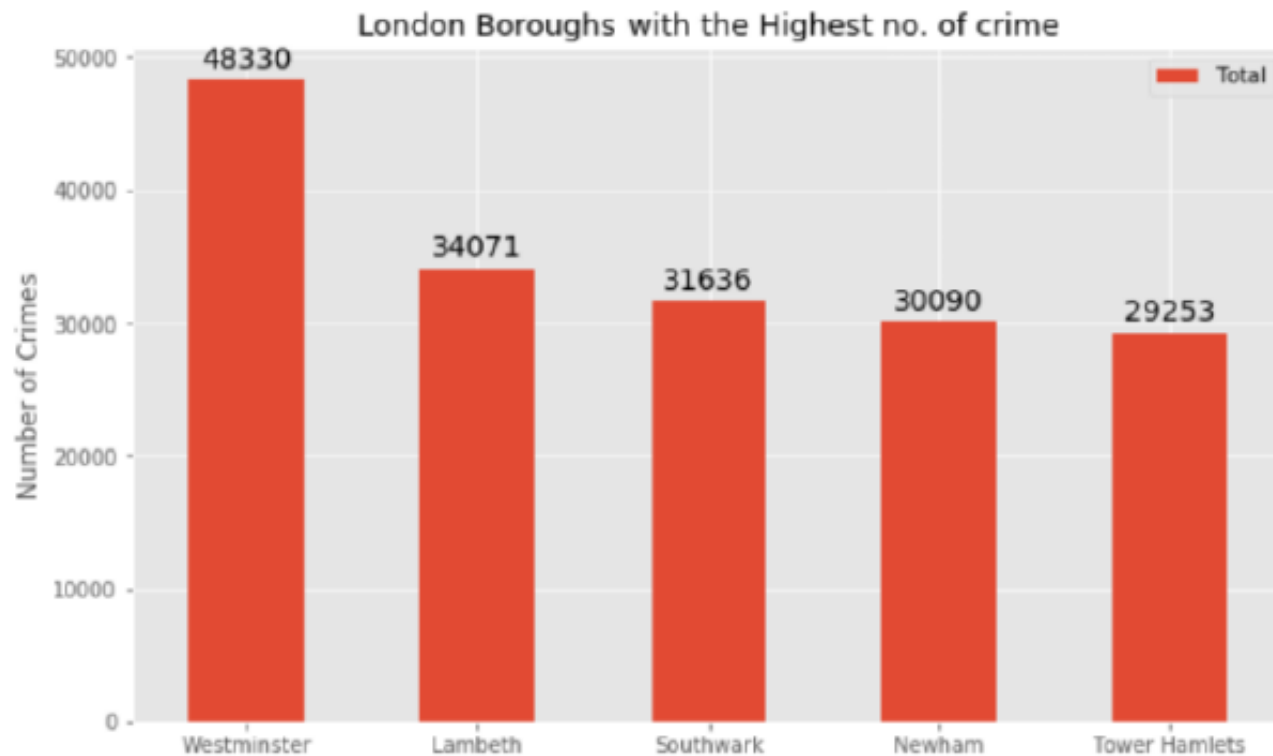
- The data cleaning process for each of the three sources of data was done separately.
 1. London Crime Data: Only the crimes during 2016 were selected. The major categories of crimes are pivoted to get the total crimes per the boroughs for each major category
 2. Data scraped from Wikipedia: This was done using the **Beautiful Soap** library in python.
 3. The two data sets are merged on the Borough names to form a new data set. The purpose of this data set is to visualize the crime rates in each borough and identify the borough with the least crimes recorded during the year 2016.

Methodology

Statistical Summary of Crimes				
ugs	Other Notifiable Offences	Robbery	Theft and Handling	Violenc
000	33.000000	33.000000	33.000000	
121	479.060606	682.666667	8913.121212	
416	223.298698	441.425366	4620.565054	
000	6.000000	4.000000	129.000000	
000	378.000000	377.000000	5919.000000	
000	490.000000	599.000000	8925.000000	
000	551.000000	936.000000	10789.000000	
000	1305.000000	1822.000000	27520.000000	

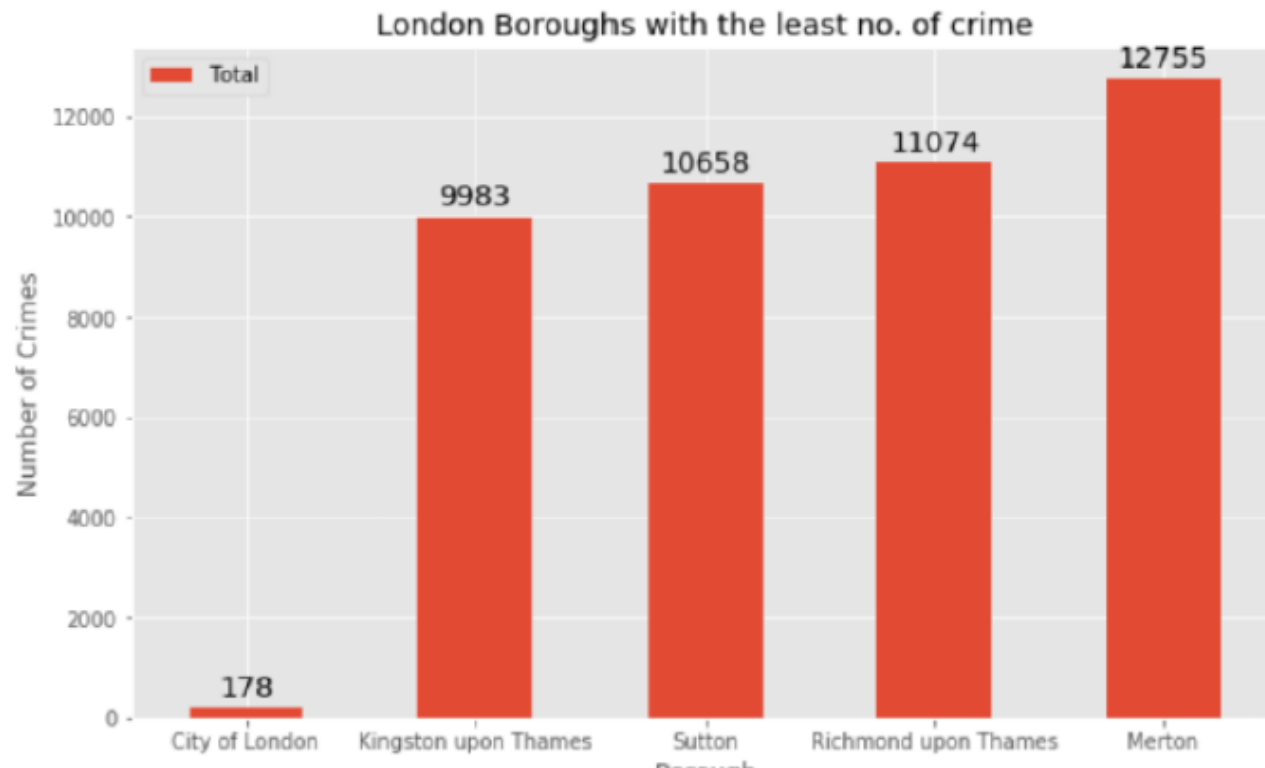
- Exploratory Data Analysis
 - The count for each of the major categories of crime returns the value 33 which is the number of London boroughs. 'Theft and Handling' is the highest reported crime during the year 2016 followed by 'Violence against person'.

Statistical Summary of Crimes



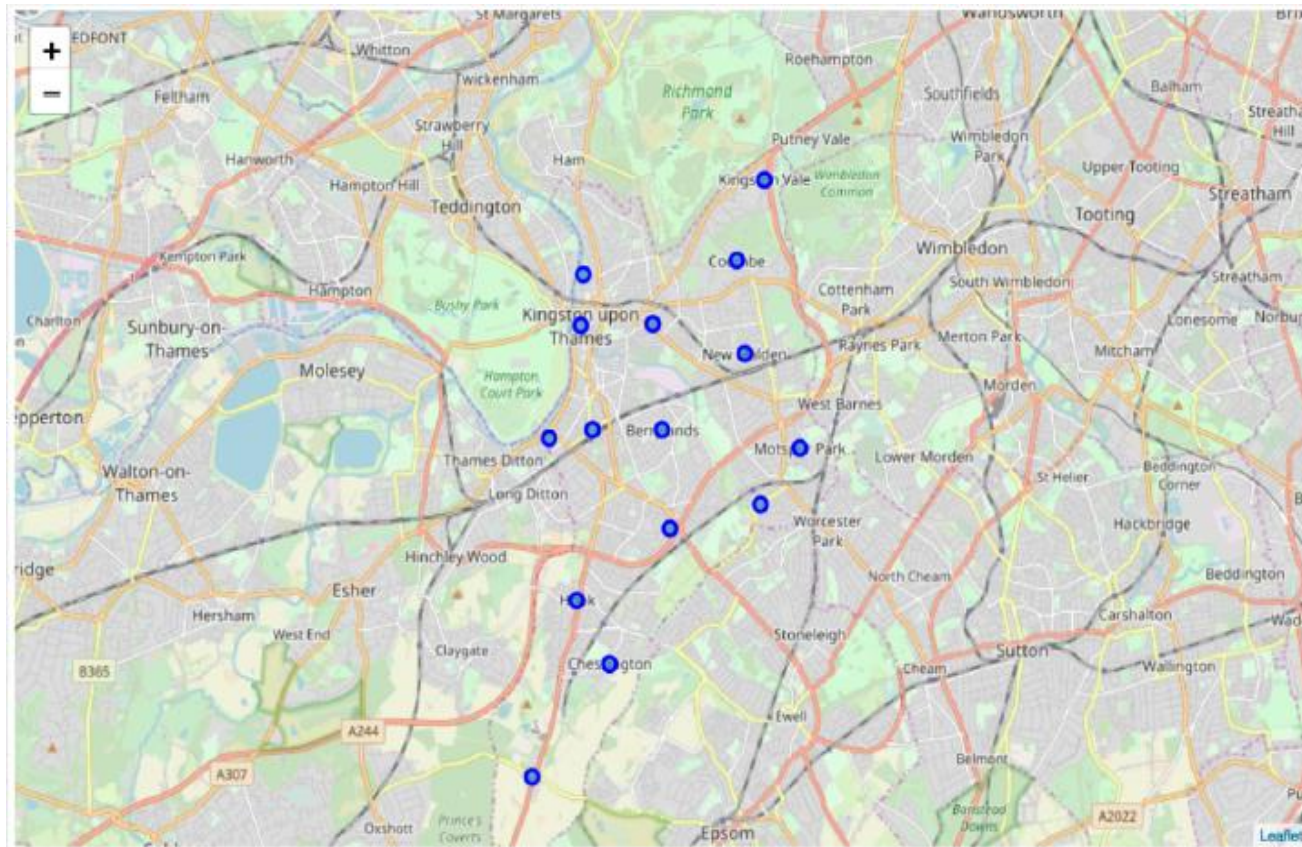
- Comparing five boroughs with the highest crime rate during the year 2016, it is evident that Westminster has the highest crimes recorded followed by Lambeth, Southwark, Newham and Tower Hamlets.
- Westminster has a significantly higher crime rate than the other 4 boroughs.

Boroughs with the Lowest Crime Rates



- Comparing five boroughs with the lowest crime rate during the year 2016, City of London has the lowest recorded crimes followed by Kingston upon Thames, Sutton, Richmond upon Thames and Merton.

Neighborhoods in Kingston upon Thames



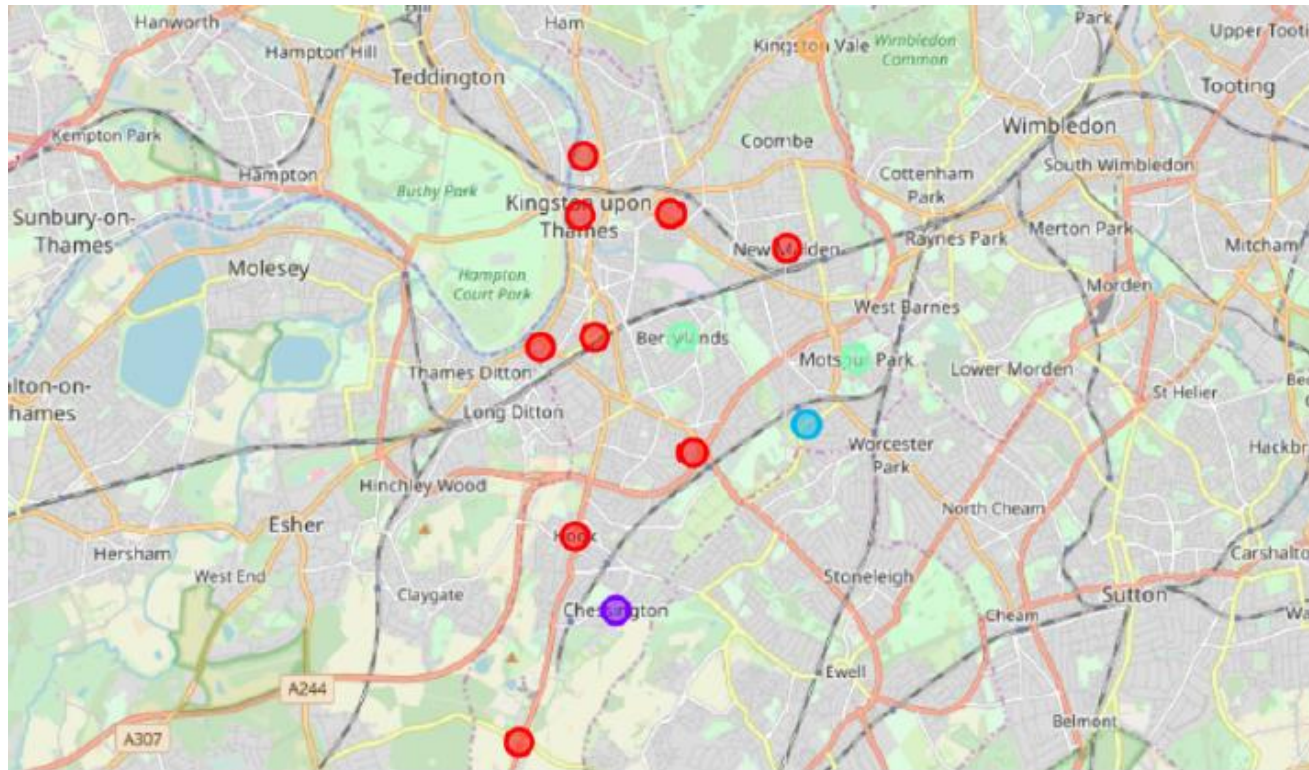
- There are 15 neighborhoods in the royal borough of Kingston upon Thames, they are visualized on a map using folium within Python.

Modelling

- Using the final data set containing the neighborhood in Kingston upon Thames along with the latitude and longitude, we can find all the values withing a 500m radius of each neighborhood by connecting to the Foursquare API.
- To help people find similar neighborhoods in the safest borough we will be clustering similar neighborhoods using K-means clustering which is a form of unsupervised machine learning algorithm that clusters data based on predefined cluster size.
- We will use a cluster size of 5 for this project.

Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Berrylands	51.393781	-0.284802	Surbiton Racket & Fitness Club	51.392676	-0.290224	Gym / Fitness Center
Berrylands	51.393781	-0.284802	Alexandra Park	51.394230	-0.281206	Park
Berrylands	51.393781	-0.284802	K2 Bus Stop	51.392302	-0.281534	Bus Stop
Berrylands	51.393781	-0.284802	Cafe Rosa	51.390175	-0.282490	Café
Canbury	51.417499	-0.305553	The Boater's Inn	51.418546	-0.305915	Pub

Results



- After running the K-means clustering algorithm we can access each cluster to see which neighborhoods were assigned to each of the five clusters.
- I visualized the clustered neighborhoods on a map using the folium library.

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
1	Canbury	Kingston upon Thames	51.417499	-0.305553	0	Pub	Café	Plaza	Fish & Chips Shop	Supermarket	Spa	Shop & Service	Park
4	Hook	Kingston upon Thames	51.367898	-0.307145	0	Bakery	Convenience Store	Indian Restaurant	Fish & Chips Shop	Wine Shop	Food	Electronics Store	Farmers Market
5	Kingston upon Thames	Kingston upon Thames	51.409627	-0.306262	0	Coffee Shop	Café	Burger Joint	Sushi Restaurant	Pub	Record Shop	Cosmetics Shop	Market
7	Malden Rushett	Kingston upon Thames	51.341052	-0.319076	0	Convenience Store	Pub	Garden Center	Restaurant	Fast Food Restaurant	Discount Store	Dry Cleaner	Electronics Store
9	New Malden	Kingston upon Thames	51.405335	-0.263407	0	Gastropub	Gym	Sushi Restaurant	Supermarket	Korean Restaurant	Indian Restaurant	Fish & Chips Shop	Dry Cleaner
10	Norbiton	Kingston upon Thames	51.409999	-0.287396	0	Indian Restaurant	Pub	Food	Italian Restaurant	Platform	Grocery Store	Farmers Market	Dry Cleaner
12	Seething Wells	Kingston upon Thames	51.392642	-0.314366	0	Indian Restaurant	Coffee Shop	Italian Restaurant	Pub	Café	Wine Shop	Fast Food Restaurant	Chinese Restaurant
13	Surbiton	Kingston upon Thames	51.393756	-0.303310	0	Coffee Shop	Pub	Supermarket	Breakfast Spot	Grocery Store	Gastropub	French Restaurant	Train Station
14	Tolworth	Kingston upon Thames	51.378876	-0.282860	0	Grocery Store	Pharmacy	Furniture / Home Store	Train Station	Pizza Place	Discount Store	Coffee Shop	Bus Stop

- The first cluster is the bigger cluster with 9/15 neighborhoods in the borough Kingston upon Thames.
- Upon closer examination we can see that the most common venues in these neighborhoods are Restaurants, Pubs, Café, Supermarkets and stores.

Cluster 2

- The second cluster has one neighborhood which consists of venues such as Restaurants, Golf courses and wine shops.

	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
2	Chessington	Kingston upon Thames	51.358336	-0.298622	1	Fast Food Restaurant	Wine Shop	Golf Course	German Restaurant	Gastropub	Garden Center	Furniture / Home Store	Fried Chicken Joint	French Restaurant

Cluster 3

- The third cluster has one neighborhood which consists of venues such as Train stations, Restaurants and Furniture shops.

	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	Train Station	Pub	Food	Gastropub	Garden Center	Furniture / Home Store	Fried Chicken Joint	French Restaurant	Deli / Bodega

Cluster 4

- The fourth cluster has two neighborhoods in it, these neighborhoods have common venues such as parks, Gym/Fitness centers, Bus stops, Restaurants, electronic stores and soccer fields etc.

	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

Cluster 5

- The fifth cluster has one neighborhood which consists of venues such as Grocery shops, Bars, Restaurants, Furniture shops, Department stores.

	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
6	Kingston Vale	Kingston upon Thames	51.43185	-0.258138	4	Grocery Store	Bar	Italian Restaurant	Soccer Field	Garden Center	Furniture / Home Store	Fried Chicken Joint	French Restaurant	Department Store

Discussion

- The aim of this project is to help people who want to relocate to the safest borough in London, expats can choose 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.
- The preference of venues may vary from person to person, they can select a neighborhood based on ones priorities.

Conclusion

- This projects 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 when it is available i.e. to find out more about places before moving into a neighborhood.
- I would like to thank my mother for giving me the opportunity to take part in this IBM Data Science Professional course.
- Thank you for your time.