

A long-exposure photograph of the Tower Bridge in London at night. The bridge's two towers are brightly lit with warm yellow lights, and the suspension cables are illuminated with blue light. The bridge spans the River Thames, and the city lights of London are visible in the background. The water in the foreground is calm, reflecting the bridge's lights and the city skyline.

The Battle of Neighborhoods in London

Project of Applied Data Science Capstone

Coursera – IBM Data Science

Pedro Pereira

2020

Project Outline

- Introduction
- Data Description
- Methodology
- Results
- Neighborhood Recommendation
- Conclusions and further work

Introduction

Where is the most suitable/ safest neighborhood in London?



- London is the capital and largest city of England and United Kingdom (UK), as well as one of most important cities in western part of Europe, with a total population of around 9 million people distributed by their 33 boroughs (including the City of London).
- It provides innumerable offer, quality of live, a diverse range of people and cultures, etc., however London is also a city with a **vast criminal record!**
- The main goal of this project is to answer the question applying data science concepts and tools based on criminality/ safety records and common venues of each neighborhood;
- This business problem is aimed to everybody who wants to rent/ buy a new house in London district or any start up company who wants to move their headquarters to a safe and comfortable neighborhood of London.

Data Description

Main Dataset: London Criminal Records

	Isao_code	borough	major_category	minor_category	value	year	month
0	E01004177	Sutton	Theft and Handling	Theft/Taking of Pedal Cycle	1	2016	8
1	E01000086	Barking and Dagenham	Theft and Handling	Other Theft Person	1	2009	5
2	E01001301	Ealing	Theft and Handling	Other Theft Person	2	2012	1
3	E01001794	Hackney	Violence Against the Person	Harassment	1	2013	2
4	E01000733	Bromley	Criminal Damage	Criminal Damage To Motor Vehicle	1	2016	4

(source: <https://www.kaggle.com/jboysen/london-crime>)

Variables of main dataset:

- **Isao_code** (code for Lower Super Output Area in Greater London (Isao)).
- **borough** (Name of London boroughs).
- **major_category** (Major categorization of crimes).
- **minor_category** (Minor categorization of crimes according to major category).
- **value** (Number of crimes monthly reported in given borough).
- **year** (Year of reported crimes, Jan/2008 - Dez/2016).
- **month** (Month of reported crimes, Jan - Dez (1-12)).

Additional Dataset: List of London Boroughs

	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	Barnet House, 2 Bristol Avenue, Colindale	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

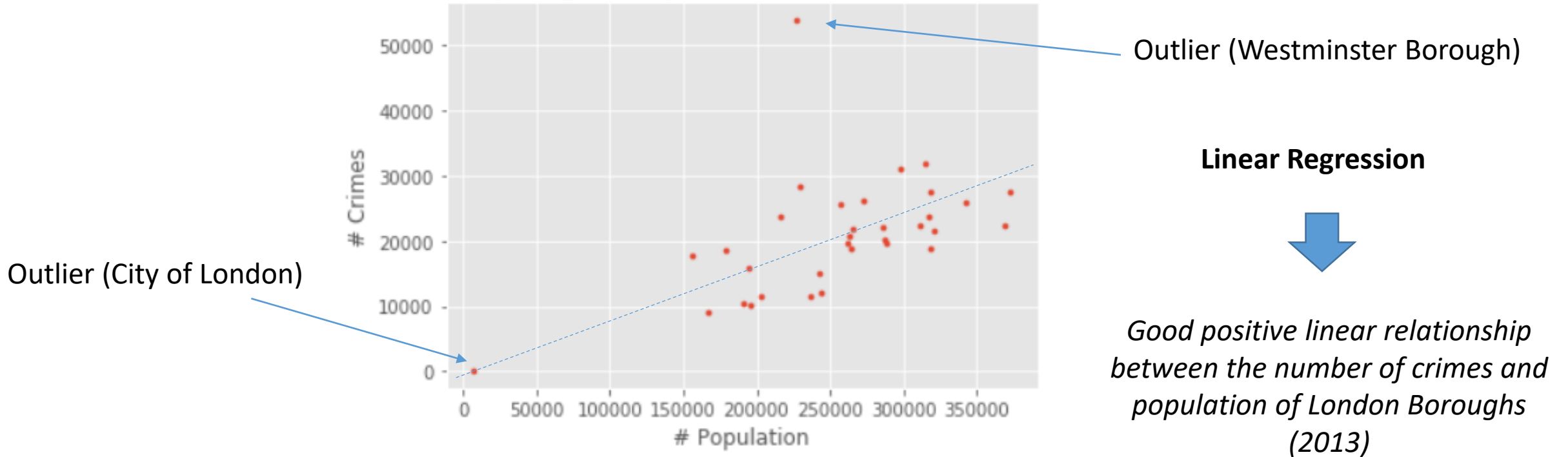
(source: https://en.wikipedia.org/wiki/List_of_London_boroughs)

Methodology

- Exploratory Data Analysis of London Criminality
 - Descriptive Analysis (histograms, scatter plots, line plots) of number of crimes, types of crimes and combination of both for safer boroughs
- Segmenting Neighborhoods of the safest borough in London
 - Application of geopy library to retrieve the geographical coordinates (latitude and longitude) of each neighborhood
 - Application of Foursquare API to explore the most common venues existing in each neighborhood area
- Clustering Neighborhoods of the safest borough in London
 - Application of K-means clustering algorithm to find (group) similar neighborhoods (through common venues)
 - Analysis of clustering output to determine the pros and cons of each cluster to select the best neighborhood

Results: Comparison between Crimes and Population

Bi-distribution comparing the Population number vs. Crimes number in 2013



- Westminster and City of London are clearly the outlier boroughs of London within this dataset;
- Although City of London is an outlier, it is visible that follows the (possible) linear regression line of the data.

Results: Numbers and Types of Crimes (2008-2016)

Total Crimes: Borough

borough	
Barking and Dagenham	149447
Barnet	212191
Bexley	114136
Brent	227551
Bromley	184349
Camden	275147
City of London	780
Croydon	260294
Ealing	251562
Enfield	193880
Greenwich	181568
Hackney	217119
Hammersmith and Fulham	185259
Haringey	213272
Harrow	116848
Havering	138947
Hillingdon	209680
Hounslow	186772
Islington	230286
Kensington and Chelsea	171981
Kingston upon Thames	89306
Lambeth	292178
Lewisham	215137
Merton	115654
Newham	262024
Redbridge	183562
Richmond upon Thames	96771
Southwark	278809
Sutton	100987
Tower Hamlets	228613
Waltham Forest	203879
Wandsworth	204741
Westminster	455028
Name: value, dtype: int64	

Crimes Type: Major Category

major_category	
Burglary	754293
Criminal Damage	630938
Drugs	470765
Fraud or Forgery	5325
Other Notifiable Offences	106349
Robbery	258873
Sexual Offences	1273
Theft and Handling	2661861
Violence Against the Person	1558081
Name: value, dtype: int64	

Crimes Type: Minor Category

minor_category	
Assault with Injury	451001
Burglary in Other Buildings	263011
Burglary in a Dwelling	491282
Business Property	21295
Common Assault	413690
Counted per Victim	3840
Criminal Damage To Dwelling	154116
Criminal Damage To Motor Vehicle	265463
Criminal Damage To Other Building	66003
Drug Trafficking	35819
Going Equipped	5530
Handling Stolen Goods	16100
Harassment	458124
Motor Vehicle Interference & Tampering	56224
Murder	949
Offensive Weapon	37983
Other Criminal Damage	145356
Other Drugs	2998
Other Fraud & Forgery	1485
Other Notifiable	100819
Other Sexual	1005
Other Theft	980085
Other Theft Person	308842
Other violence	70778
Personal Property	237578
Possession Of Drugs	431948
Rape	268
Theft From Motor Vehicle	569956
Theft From Shops	345142
Theft/Taking Of Motor Vehicle	216538
Theft/Taking of Pedal Cycle	168974
Wounding/GBH	125556
Name: value, dtype: int64	

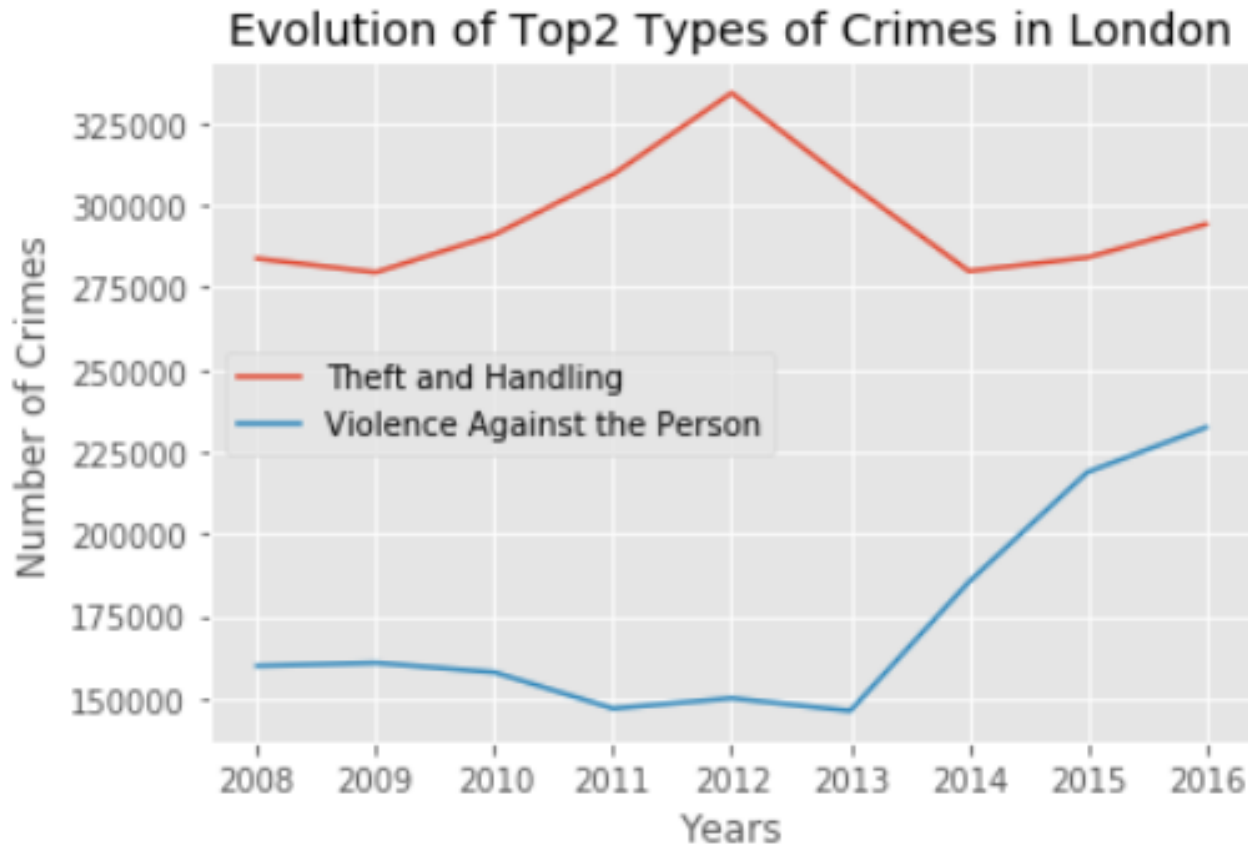
Total Crimes: Year

year	
2008	738641
2009	717214
2010	715324
2011	724915
2012	737329
2013	686407
2014	680183
2015	711624
2016	736121
Name: value, dtype: int64	

Total Crimes: Month

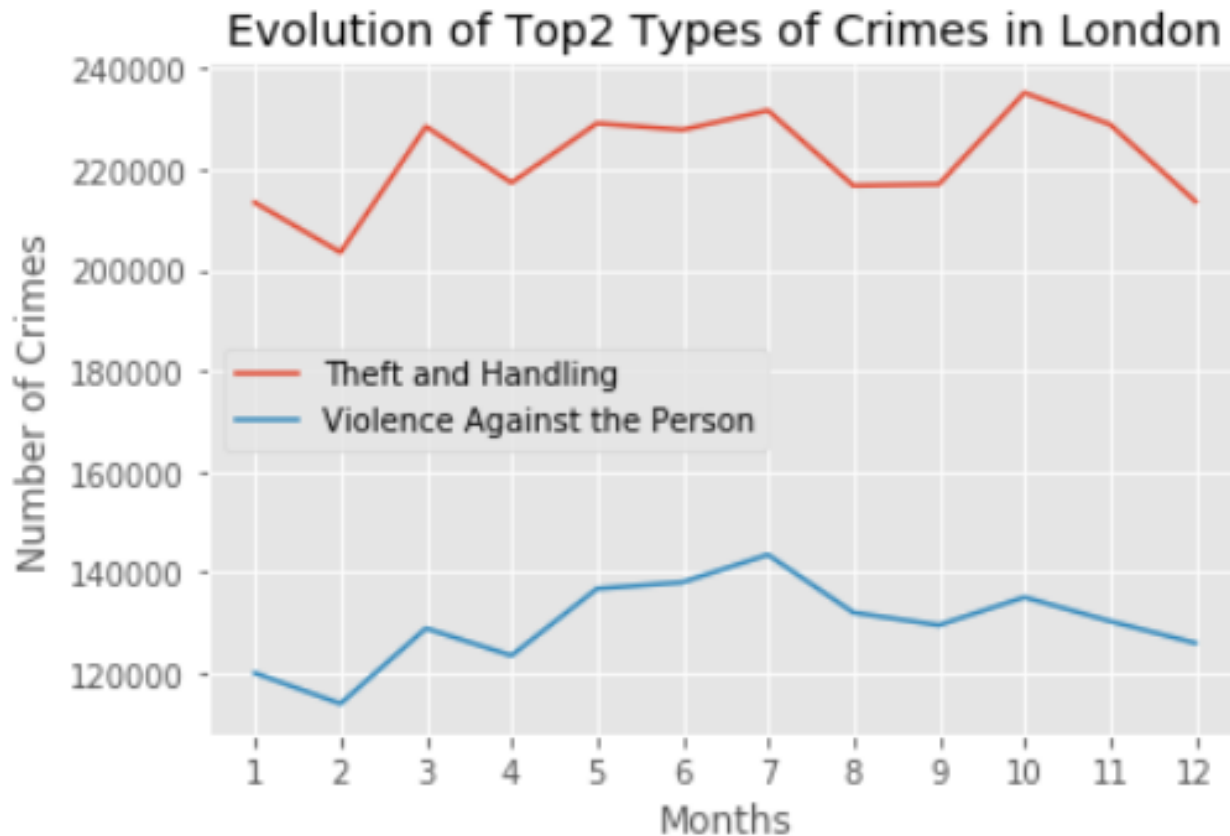
month	
1	530247
2	500537
3	551795
4	520648
5	552704
6	549041
7	563076
8	532784
9	522377
10	558252
11	549671
12	516626
Name: value, dtype: int64	

Results: Annual Evolution of Top2 Crimes in London



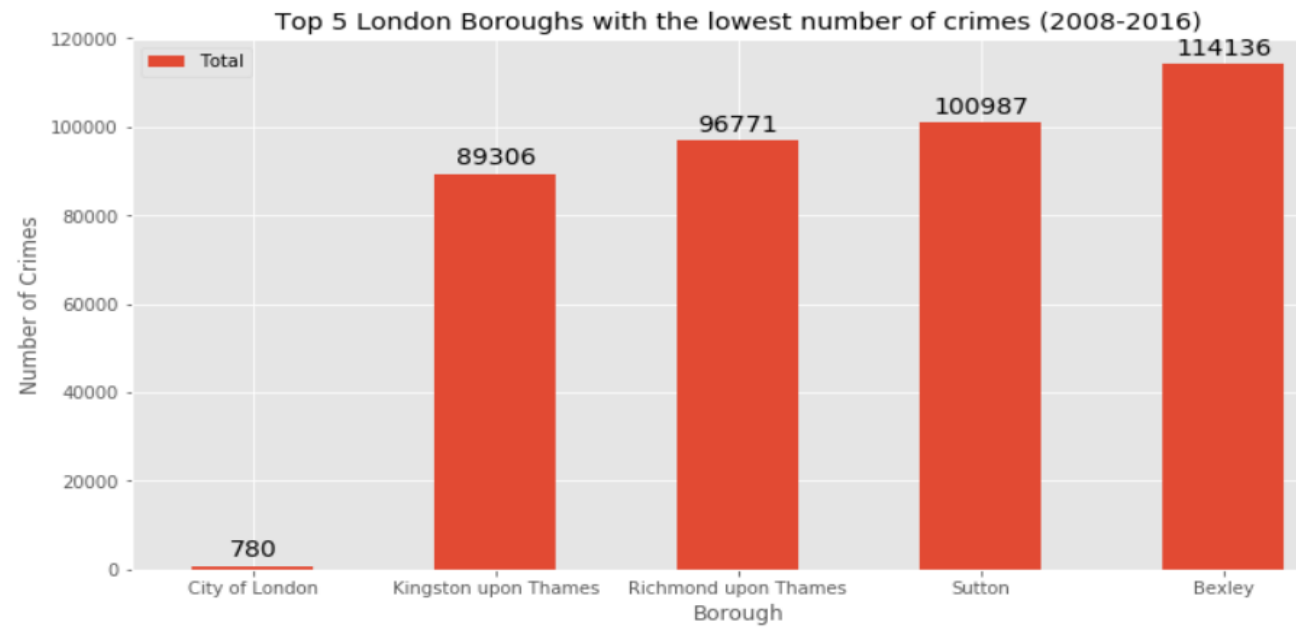
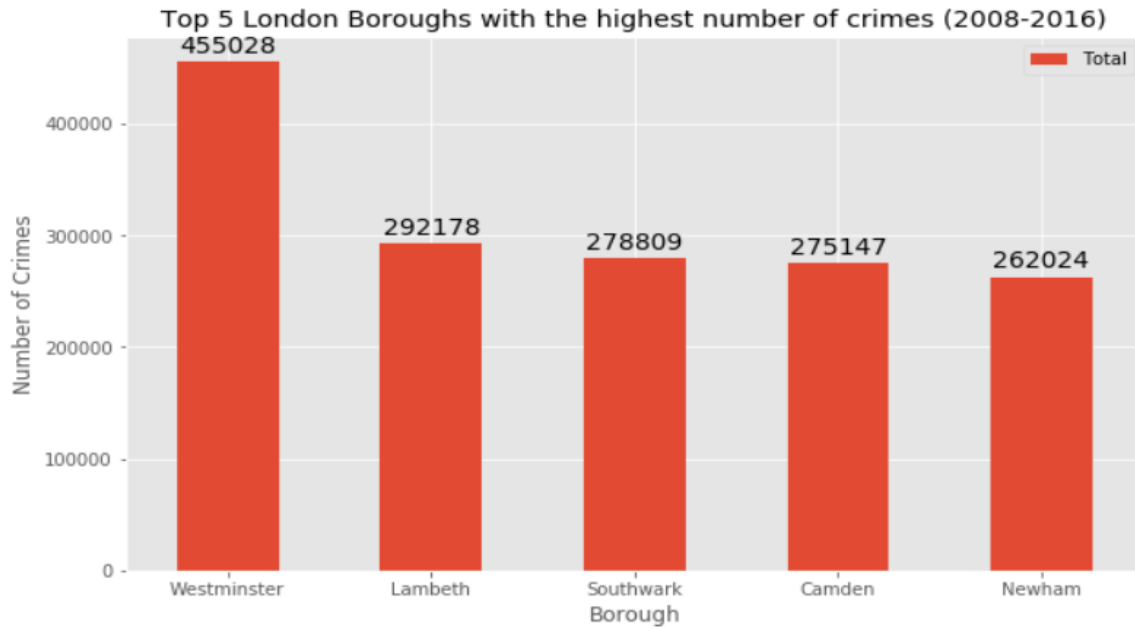
- General behavior is quite different between both crime types;
- Highest number of reports happened in 2012 for “Theft and Handling”;
- In spite of “Violence Against the Person” crime type is the second most often in London, its behavior was stable between 2008 and 2013, significantly increasing afterwards.

Results: Monthly Evolution of Top2 Crimes in London



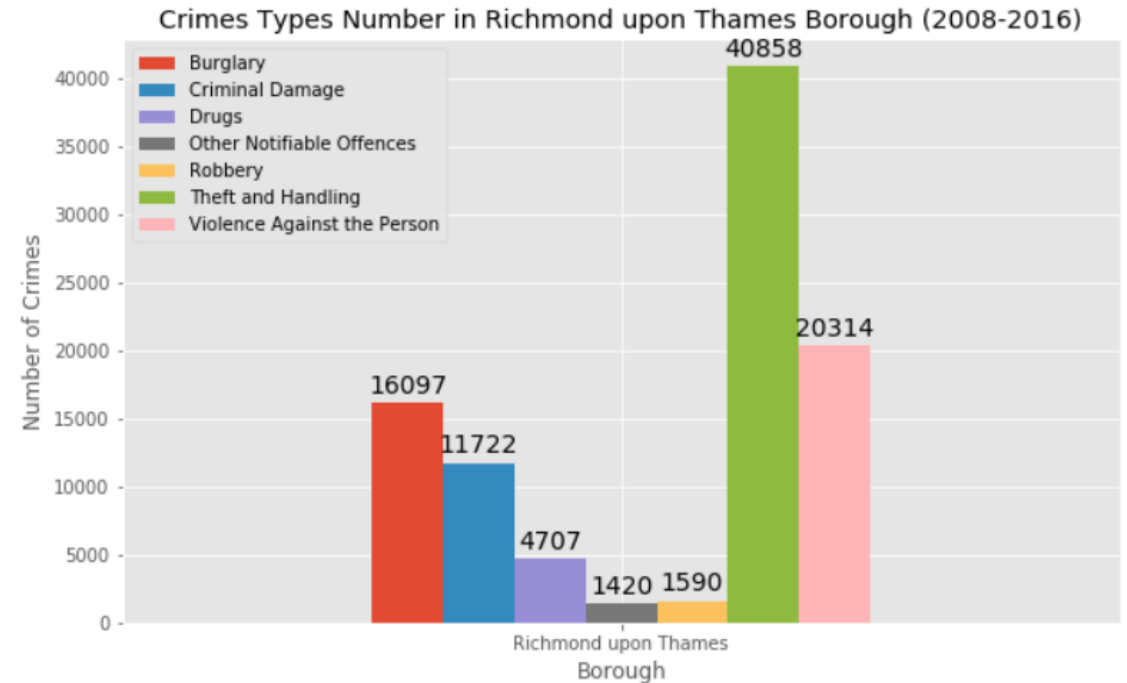
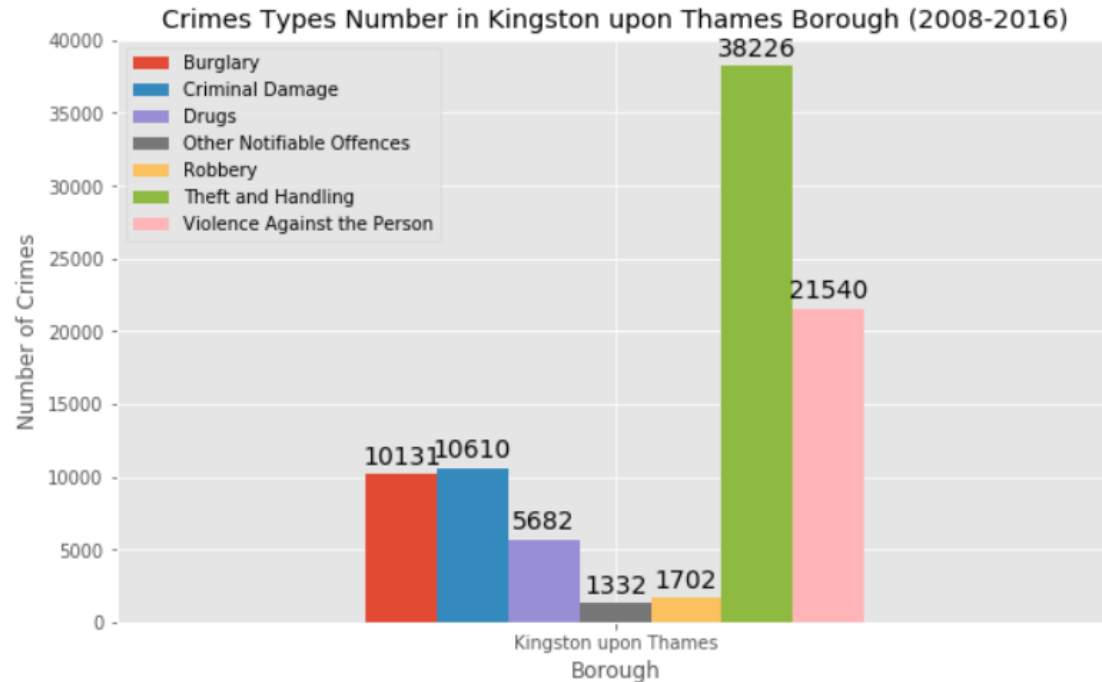
- General evolution behavior between “Theft and Handling” and “Violence Against the Person” crimes is similar over the months;
- Between 2008 and 2016:
 - the most critical months were March, July and October for both crime types;
 - February and December were the “safest” months with less reported cases

Results: Total Crimes in London Boroughs



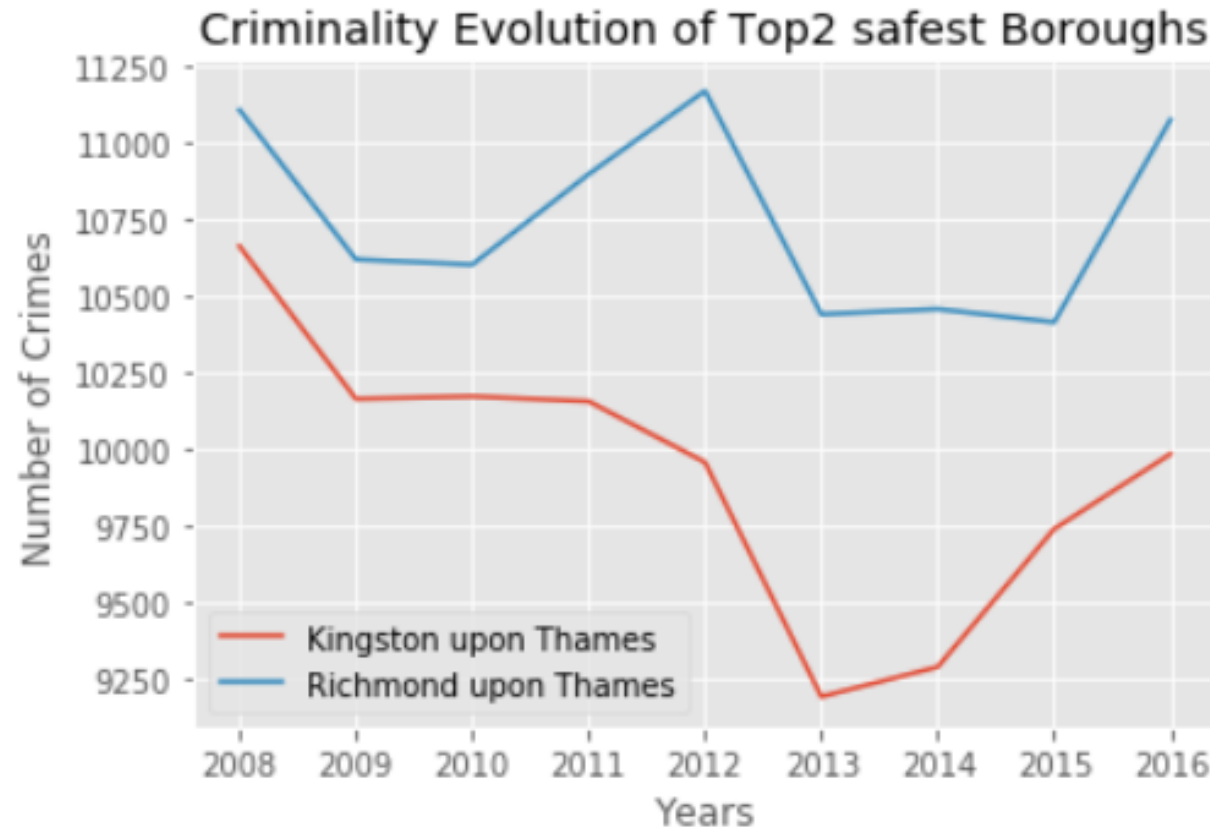
- “Westminster” Borough presents the higher number of reported crimes in London, therefore it’s the most **dangerous**;
- “City of London” is the **safest** Borough in London with 780 reported crimes between 2008-2016;
- Since “City of London” is not trully considered a London Borough, next results are further analysed between “Kingston upon Thames” and “Richmond upon Thames” to select the safest Borough in London.

Results: Total Crime Types in Top2 safest Boroughs



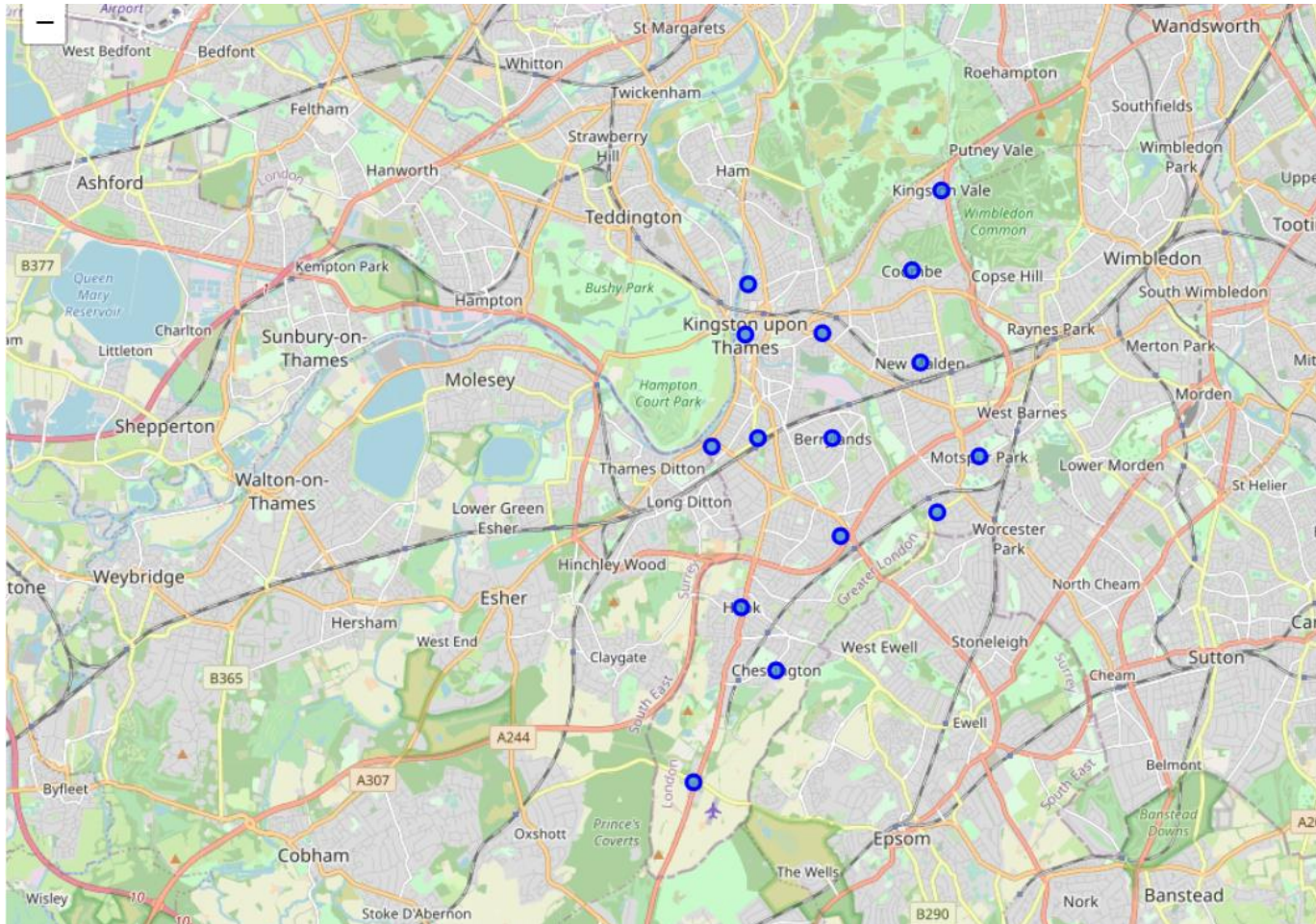
- General behavior of Crimes Types, within Major Category, is quite similar between both analysed Boroughs (although frequency number is different);
- Clearly the crime types “Theft and Handling” and “Violence Against the Person” happen more often;
- Similarly between both Boroughs, crime types classified as “Robbery” and “Other Notifiable Offences” are the less reported.

Results: Evolution of Total Crimes in Top2 safest Boroughs



- Between 2008-2009 and 2015-2016 the evolution behavior is quite similar;
- Critical years of criminality maximums were 2008, 2012 and 2016, for “Richmond upon Thames”, and mostly 2008 for “Kingston upon Thames”;
- “Richmond upon Thames” registered a peak in crimes between 2010 and 2013;
- “Kingston upon Thames” shows a significant decrease in crimes between 2012 and 2013.

Results: Segmenting Neighborhoods of Kingston upon Thames



	Neighborhood	Borough	Latitude	Longitude
0	Berrylands	Kingston upon Thames	51.393781	-0.284802
1	Canbury	Kingston upon Thames	51.417499	-0.305553
2	Chessington	Kingston upon Thames	51.358336	-0.298622
3	Coombe	Kingston upon Thames	51.419450	-0.265398
4	Hook	Kingston upon Thames	51.367898	-0.307145
5	Kingston upon Thames	Kingston upon Thames	51.409627	-0.306262
6	Kingston Vale	Kingston upon Thames	51.431850	-0.258138
7	Malden Rushett	Kingston upon Thames	51.341052	-0.319076
8	Motspur Park	Kingston upon Thames	51.390985	-0.248898
9	New Malden	Kingston upon Thames	51.405335	-0.263407
10	Norbiton	Kingston upon Thames	51.409999	-0.287396
11	Old Malden	Kingston upon Thames	51.382484	-0.259090
12	Seething Wells	Kingston upon Thames	51.392642	-0.314366
13	Surbiton	Kingston upon Thames	51.393756	-0.303310
14	Tolworth	Kingston upon Thames	51.378876	-0.282860

Results: Clustering Neighborhoods of Kingston upon Thames

K-means clustering algorithm
(Search radius = 750 meters)

Cluster 0:

- Berrylands
- Chessington

Cluster 1:

- Coombe

Cluster 3:

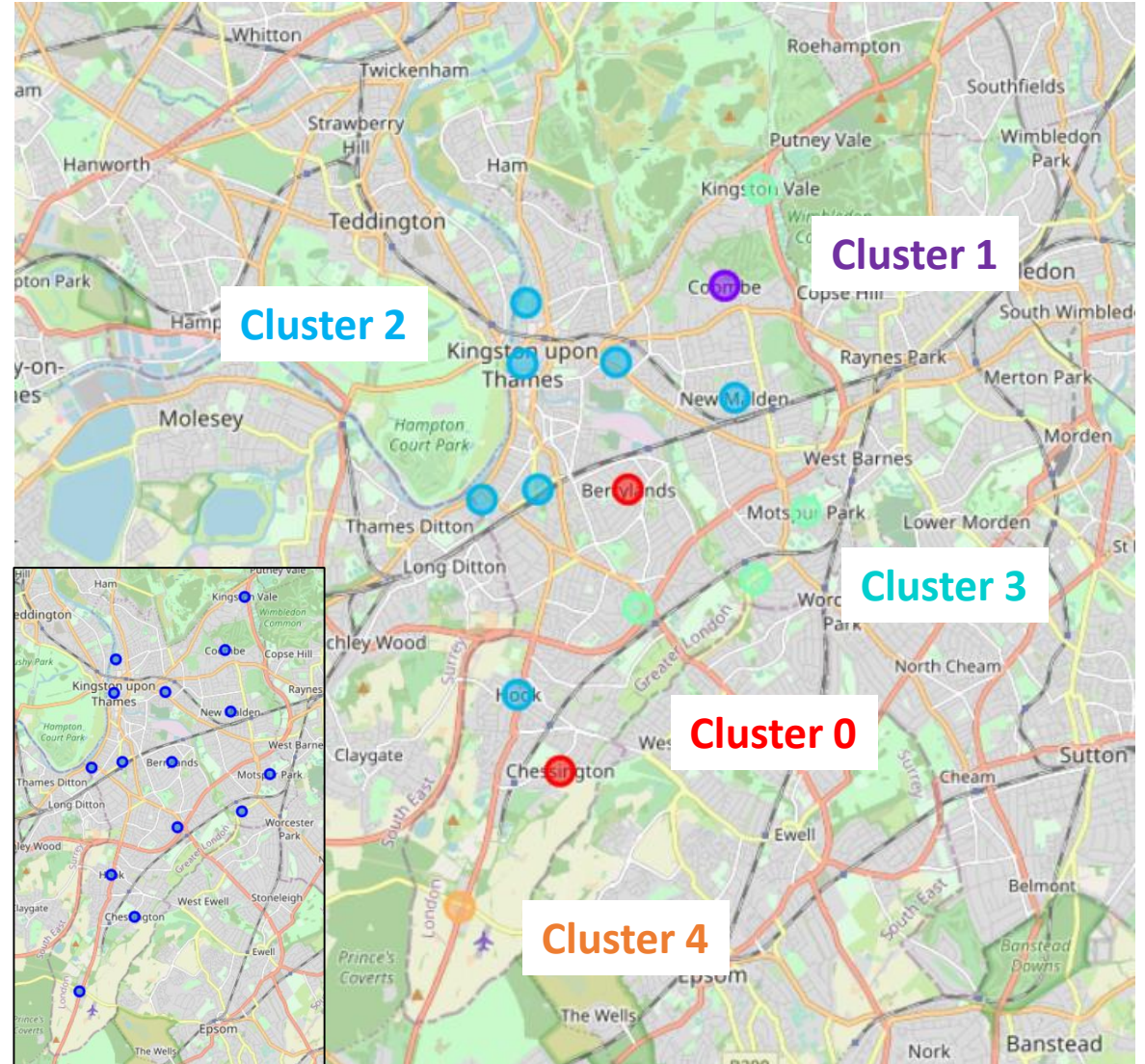
- Kingston Vale
- Motspur Park
- Old Malden
- Tolworth

Cluster 2:

- Canbury
- Hook
- Kingston upon Thames
- New Malden
- Norbiton
- Seething Wells
- Surbiton

Cluster 4:

- Malden Rushett



Results: Clusters 0 and 1

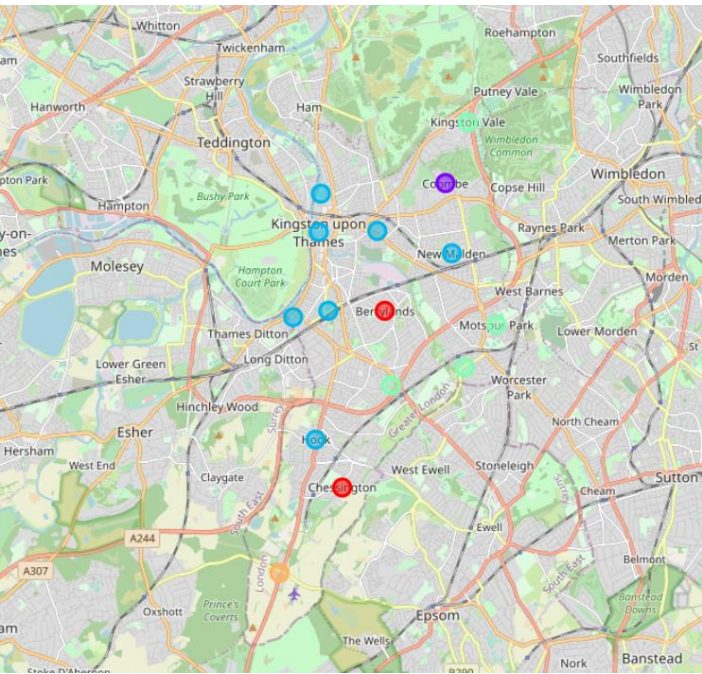
Most common venues:

Cluster 0:

- Pub
- Train Station
- Bus Stop
- Coffee Shop
- Golf Course

Cluster 1:

- Golf Course
- Hotel
- Garden
- Spa
- Food



Cluster 0 (red points)
Cluster 1 (purple point)

Cluster 0

	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	10th Most Common Venue
0	Berrylands	Kingston upon Thames	51.393781	-0.284802	0	Pub	Bus Stop	Coffee Shop	Park	Platform	Gym / Fitness Center	Train Station	Farmers Market	Department Store	Discount Store
2	Chessington	Kingston upon Thames	51.358336	-0.298622	0	Train Station	Golf Course	Breakfast Spot	Convenience Store	Fast Food Restaurant	Platform	Garden	Furniture / Home Store	Fried Chicken Joint	French Restaurant

Cluster 1

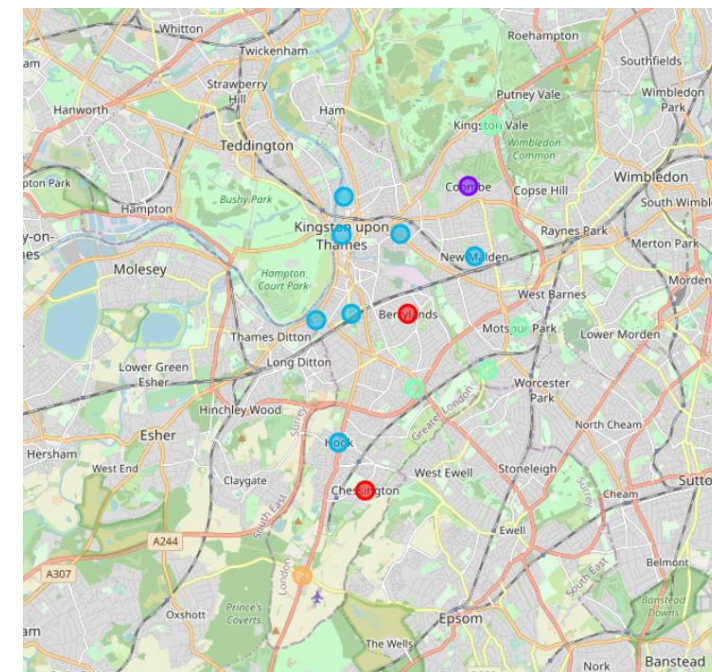
	Neighborhood	Borough	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
3	Coombe	Kingston upon Thames	51.41945	-0.265398	1	Golf Course	Hotel	Garden	Spa	Food	Discount Store	Donut Shop	Electronics Store	Farmers Market	Fast Food Restaurant

Results: Cluster 2

Most common venues:

- Café
- Breakfast Spot
- Coffee Shop
- Supermarket
- Restaurants (Korean, Italian, Indian)

Clusters in blue (map)



	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	10th Most Common Venue
1	Canbury	Kingston upon Thames	51.417499	-0.305553	2	Café	Coffee Shop	Pub	Hotel	Italian Restaurant	Supermarket	Plaza	Turkish Restaurant	Fish & Chips Shop	Indian Restaurant
4	Hook	Kingston upon Thames	51.367898	-0.307145	2	Breakfast Spot	Indian Restaurant	Convenience Store	Park	Fast Food Restaurant	Fish & Chips Shop	Supermarket	Platform	Bakery	Grocery Store
5	Kingston upon Thames	Kingston upon Thames	51.409627	-0.306262	2	Café	Coffee Shop	Pub	Burger Joint	Sushi Restaurant	Mexican Restaurant	Department Store	Cosmetics Shop	Clothing Store	Portuguese Restaurant
9	New Malden	Kingston upon Thames	51.405335	-0.263407	2	Korean Restaurant	Supermarket	Indian Restaurant	Fast Food Restaurant	Clothing Store	Café	Used Auto Dealership	Karaoke Bar	Department Store	Gastropub
10	Norbiton	Kingston upon Thames	51.409999	-0.287396	2	Italian Restaurant	Gastropub	Indian Restaurant	Pub	Fried Chicken Joint	Gym / Fitness Center	Japanese Restaurant	Pharmacy	Pizza Place	Coffee Shop
12	Seething Wells	Kingston upon Thames	51.392642	-0.314366	2	Coffee Shop	Pub	Indian Restaurant	Restaurant	Gastropub	French Restaurant	Golf Course	Grocery Store	Gym / Fitness Center	Fish & Chips Shop
13	Surbiton	Kingston upon Thames	51.393756	-0.303310	2	Coffee Shop	Pub	Grocery Store	Restaurant	French Restaurant	Fish & Chips Shop	Indian Restaurant	Italian Restaurant	Farmers Market	Deli / Bodega

Results: Clusters 3 and 4

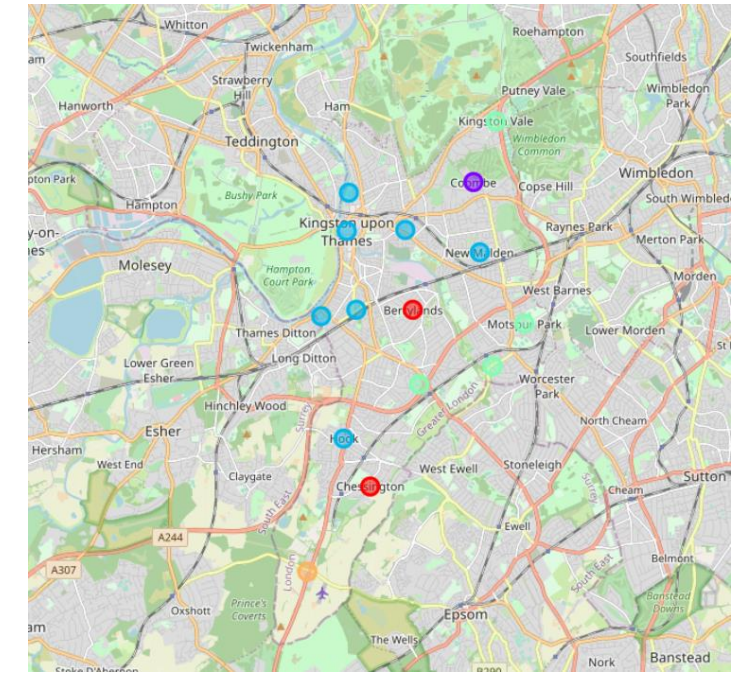
Most common venues:

Cluster 3:

- Coffee Shop
- Gym/ Fitness Center
- Grocery Store
- Tennis Court
- Soccer Field

Cluster 4:

- Theme Park Ride/ Attraction
- Pub
- Restaurant
- Garden Center
- Hotel



Cluster 3 (light blue points)

Cluster 4 (orange point)

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	10th Most Common Venue
6	Kingston Vale	Kingston upon Thames	51.431850	-0.258138	3	Coffee Shop	Stables	Outdoors & Recreation	Sandwich Place	Soccer Field	Bus Stop	Grocery Store	Bar	Bistro	Food Truck
8	Motspur Park	Kingston upon Thames	51.390985	-0.248898	3	Gym	Tennis Court	Park	Soccer Field	Bus Stop	Japanese Restaurant	Steakhouse	Bakery	Grocery Store	Furniture / Home Store
11	Old Malden	Kingston upon Thames	51.382484	-0.259090	3	Grocery Store	Gym / Fitness Center	Japanese Restaurant	Bakery	Steakhouse	Train Station	Gastropub	Food	Discount Store	Donut Shop
14	Tolworth	Kingston upon Thames	51.378876	-0.282860	3	Grocery Store	Soccer Field	Restaurant	Coffee Shop	Italian Restaurant	Pizza Place	Discount Store	Garden Center	Sandwich Place	Café

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	10th Most Common Venue
7	Malden Rushett	Kingston upon Thames	51.341052	-0.319076	4	Theme Park Ride / Attraction	Pub	Restaurant	Garden Center	Hotel	Grocery Store	Bar	Food Truck	Donut Shop	Electronics Store

Neighborhood Recommendation

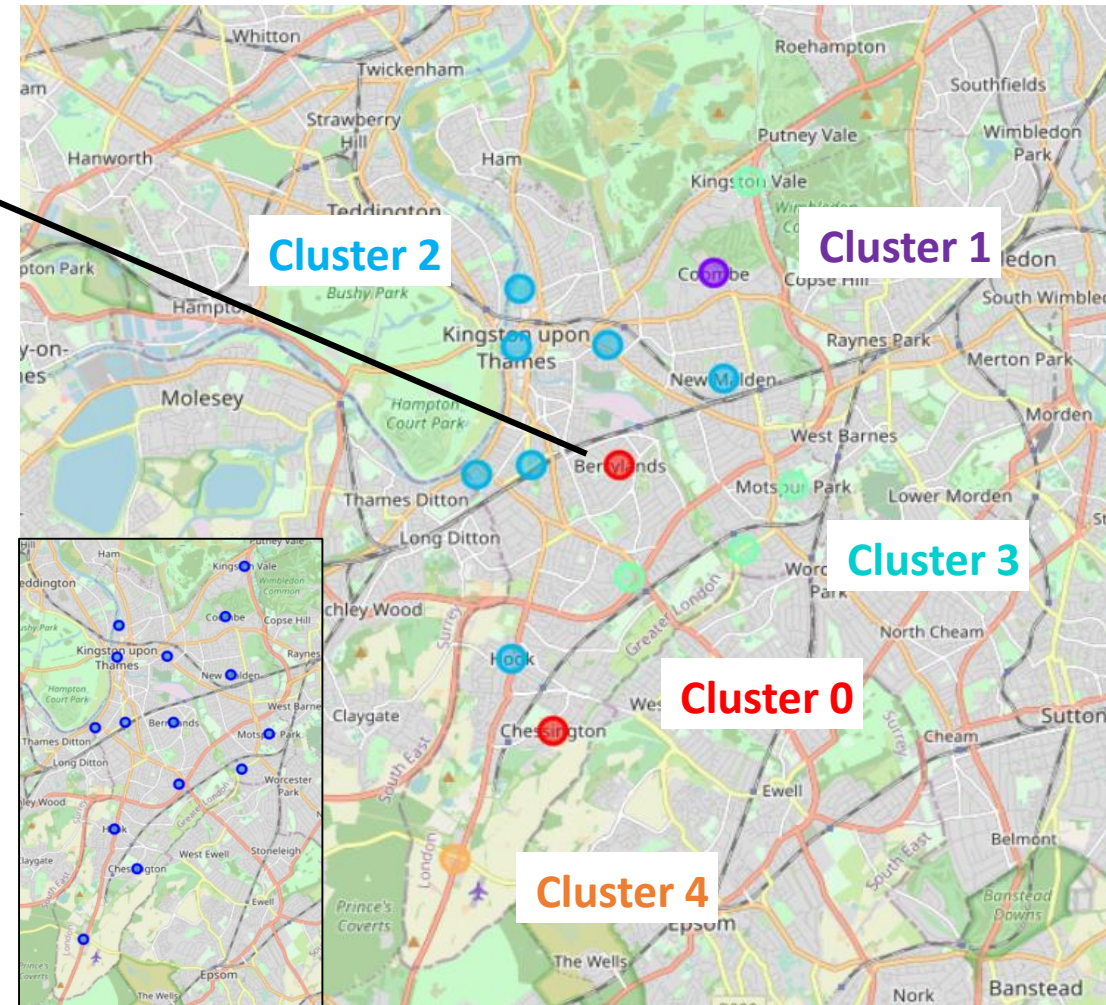
Cluster 0: Berrylands neighborhood

Common Venues (max $\leq 750\text{m}$):

- Pub
- Train Station
- Bus Stop
- Coffee Shop
- Golf Course

Common Venues from adjacent clusters 2 and 3 ($> 750\text{m}$):

- Café
- Gym/ Fitness Center
- Breakfast Spot
- Supermarket
- Grocery Store
- Restaurants (Korean, Italian, Indian)
- Tennis Court
- Soccer Field
- Golf Course
- Hotel/ Spa
- Garden



Conclusions

- In this project the necessary exploratory analysis of London crime data was done to better understand how criminality components evolved during last years and to find the safest borough (Kingston upon Thames);
- Segmenting and clustering of neighborhoods belonging to Kingston upon Thames borough was achieved using Foursquare API and K-means algorithm, allowing to explore and retrieve the most common venues around each neighborhood and cluster them based on venues similarity;
- After analysing the several clusters and respective venues, the final recommendation to everybody who is interested to find a new house, or a new work office for a start-up company, would be Berrylands neighborhood.
- As a future work of this project, the extension of this method to other London boroughs would be valuable but also the incorporation of property/ offices costs to complement this work, since only topics of safety and common venues were taken in consideration.

References

London Crime records:

<https://www.kaggle.com/jboysen/london-crime>

List of London boroughs:

https://en.wikipedia.org/wiki/List_of_London_boroughs

One hot encoding:

<https://hackernoon.com/what-is-one-hot-encoding-why-and-when-do-you-have-to-use-it-e3c6186d008f>

List of neighborhoods in Kingston upon Thames borough:

https://en.wikipedia.org/wiki/List_of_districts_in_the_Royal_Borough_of_Kingston_upon_Thames