1. Data

1.1 Data Utilized

For this project we need data about:

- Ø London Crime Rates obtained from Kaggle;
- Ø List of London boroughs; obtained and scraped from a Wikipedia page.
- Ø List of neighborhoods in the Royal Borough of Kingston upon Thames also from Wikipedia page.

1.2 Data Cleaning

ü *London Crime Data:*

Data cleaning of the three (3) data sets were done separately; in the London crime data, the latest year (2016), is only selected. The major categories of crime are swivelled to get the total crimes per the boroughs for each major category. *(See the below fig 2.1, After Pre-processing the data)*

Major_Category	Borough	Burglary	Criminal Damage	Drugs	Other Notifiable Offences	Robbery	Theft and Handling	No_of_Crimes Violence Against the Person	Total
0	Barking and Dagenham	1287	1949	919	378	534	5607	6067	16741
1	Barnet	3402	2183	906	499	464	9731	7499	24684
2	Bexley	1123	1673	646	294	209	4392	4503	12840
3	Brent	2631	2280	2096	536	919	9026	9205	26693
4	Bromley	2214	2202	728	417	369	7584	6650	20164
5	Camden	2652	1935	1493	490	899	14088	7626	29183
6	City of London	2	2	10	6	4	129	25	178
7	Croydon	2738	3219	1367	718	1139	9229	10302	28712
8	Ealing	2492	2562	1355	613	669	10040	9396	27127
9	Enfield	2541	2136	1063	492	807	8037	7409	22485
10	Greenwich	1780	2476	867	521	486	8010	8590	22730

ü *List of London Boroughs (Wikipedia):*
The London boroughs data set taken from Wikipedia is scraped using BeautifulSoup library in python. This library is used in extracting the data in tabular format as in the Wikipedia website. String manipulation is obvious after scraping, to get names of the boroughs in the exact format. Its very essential to as we will be merging the two (2) datasets together using Borough names. *(See the below fig. 2.2; 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	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

ü *Datasets Merged (London Crime Data & List of London Boroughs):*
In other to gain the necessary information we need; the two datasets are merged into one dataset *(Ld_crime)*. In other to visualize the crime rates in each borough and identify the borough with the least crimes recorded during the year 2016. (*See the below fig. 2.3; London Borough Crime)*

[38]:	Borough	Local authority	Political control	Headquarters	Area (sq mi)	Population (2013 est)[1]	Co-ordinates	Burglary	Criminal Damage	Drugs	Other Notifiable Offences	Robbery	Theft and Handling	Violence Against the Person	Total
0	Barking and Dagenham	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 /...	1287	1949	919	378	534	5607	6067	16741
1	Barnet	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 /...	3402	2183	906	499	464	9731	7499	24684
2	Bexley	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 /...	1123	1673	646	294	209	4392	4503	12840
3	Brent	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 /...	2631	2280	2096	536	919	9026	9205	26693
4	Bromley	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 /...	2214	2202	728	417	369	7584	6650	20164

ü *Safest Borough:*
The borough with the lowest crime rate, automatically is the safest borough; after we visualized the crime rate in each borough. The third set of data is created from scratch, with pandas data frame the names of the neighborhoods and the name of the borough with the latitude left blank. Coordinates of the neighborhood is to be gotten from Google Maps API geocoding to get the final dataset. *(See the below fig. 2.4-5; Safest borough)*

[47]:	Neighborhood	Borough	Latitude	Longitude
0	Berrylands	Kingston upon Thames		
1	Canbury	Kingston upon Thames		
2	Chessington	Kingston upon Thames		
3	Coombe	Kingston upon Thames		
4	Hook	Kingston upon Thames		
5	Kingston upon Thames	Kingston upon Thames		
6	Kingston Vale	Kingston upon Thames		
7	Malden Rushett	Kingston upon Thames		
8	Motspur Park	Kingston upon Thames		
9	New Malden	Kingston upon Thames		
10	Norbiton	Kingston upon Thames		

[49]:	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

The Foursquare API will be used to generate and obtain the 10 most common venues for each neighborhood, K-Mean clustering algorithm will be used finally to cluster similar neighborhoods together.