Capstone Project- The Battle of Neighbourhoods

By Karthiga

Background:

Safety is top concern when moving to new area. If you don' t feel safe in your home, you are not able to enjoy living there.

Problem :

This project aim is to select safest borough in London based on total crimes by exploring neighborhoods of borough.

Interest:

Some one who is interested to relocate to London will be interested to identify safest borough in London.

Data Acquisition and Cleaning

Data Acquisition :

The acquired for this project is combination of data from Wikipedia and Kaggle dataset and also list of neighborhoods in the royal borough of Kingston.

Data Cleaning :

- Different data sources cleaned separately.

- Only 2016 data considered for this project.

- Wikipedia data is extracted from beautiful python package.

- Merge two dataset and visualize.

- After visualizing found lowest crime recorder boroughs.

Methodology

EDA (Exploratory Data Analysis)

	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	8832.000000	27174.000000
max	3402.000000	3219.000000	2738.000000	1305.000000	1822.000000	27520.000000	10834.000000	48330.000000

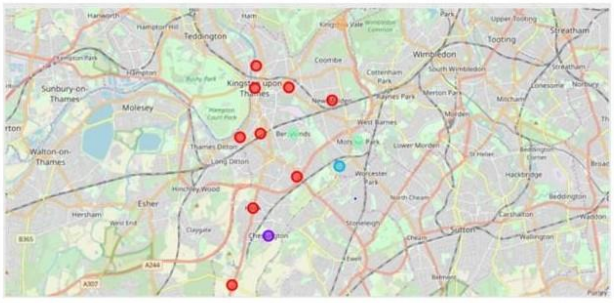
Modelling

	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	Berrylands	51.393781	-0.284802	Cafe Rosa	51.390175	-0.282490	Café
4	Canbury	51.417499	-0.305553	The Boater's Inn	51.418546	-0.305915	Pub

- Using data containing neighborhoods in Kingston along with latitude and longitude we can find all the venues using foursquare API.
- One hot coding is done on the venues data.
- We will use cluster size 5.
- We can use k-means clustering unsupervised algo.

Result

We can find K-means clustering result below.Each cluster is represented with ease of presentation



Discussion:

- The aim of this project is to help lot
- people who want to relocate to the safest borough in London, expats can chose the neighbourhoods to which they want to relocate based on the most common venues in it.
- As per visualization I feel that cluster 4 is more suitable due to common venues in that cluster.
- If person looking for train or bus stations cluster 3 & 4 are more suitable.
- The preference may varies from person to person based on priorities.

- 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.