The Battle of Neighborhoods

Introduction

- Background: To find a proper and sate place to live when coming to a new area is difficult.
- Problem: Aims to help people identify safer zone around London
- Interest: People who looking for a new place or who interested with the crime rate of London

Data acquisition and cleaning

- Data acquisition and cleaning:
  - Main data from Kaggle compititeion which is the London crime data
  - The list of London boroughs are scrapped from wiki page by using the python package.
  - Different sources of data are merged into one data frame

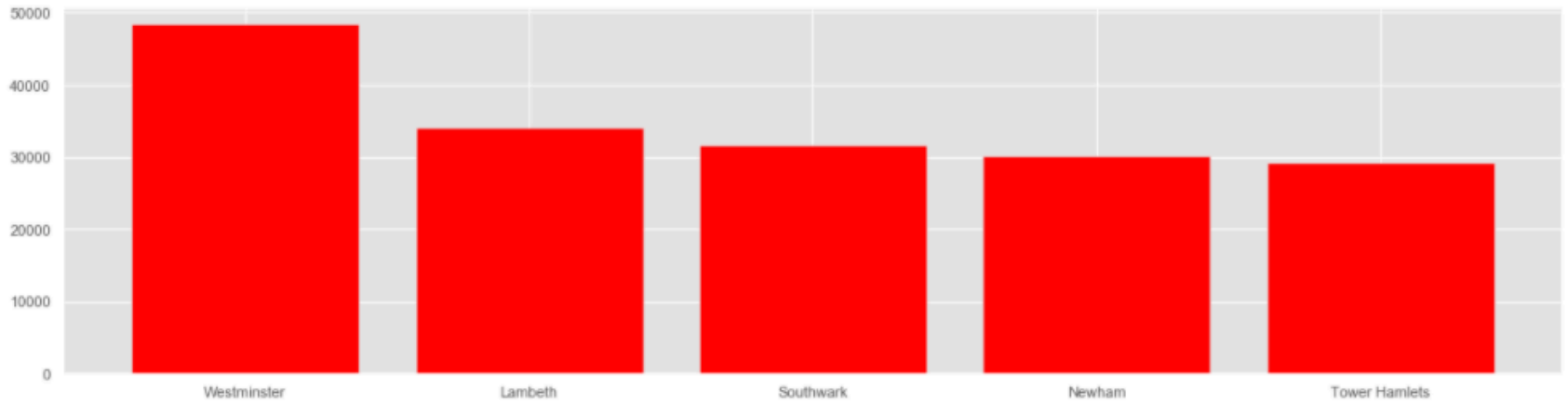
Methodology

- EDA

	Buglary	Criminal Damage	Drugs	Other Notifiable Offences	Robbery	Theft and Handling	Violence Aginst the Person	Total	Area (sq mi)	Population (2013 est)[1]
count	33.000000	33.000000	33.000000	33.000000	33.000000	33.000000	33.000000	33.000000	33.000000	32.000000
mean	2069.242424	1941.545455	1179.212121	479.060606	682.666667	8913.121212	7041.848485	22306.696970	18.408485	262777.718750
std	737.448644	625.207070	586.406416	223.298698	441.425366	4620.565054	2513.601551	8828.228749	12.645367	57346.611409
min	2.000000	2.000000	10.000000	6.000000	4.000000	129.000000	25.000000	178.000000	1.120000	155594.000000
25%	1531.000000	1650.000000	743.000000	378.000000	377.000000	5919.000000	5936.000000	16903.000000	10.360000	224047.500000
50%	2071.000000	1989.000000	1063.000000	490.000000	599.000000	8925.000000	7409.000000	22730.000000	14.520000	263697.000000
75%	2631.000000	2351.000000	1617.000000	551.000000	936.000000	10789.000000	8832.000000	27174.000000	21.780000	311447.500000
max	3402.000000	3219.000000	2738.000000	1305.000000	1822.000000	27520.000000	10834.000000	48330.000000	57.970000	372752.000000

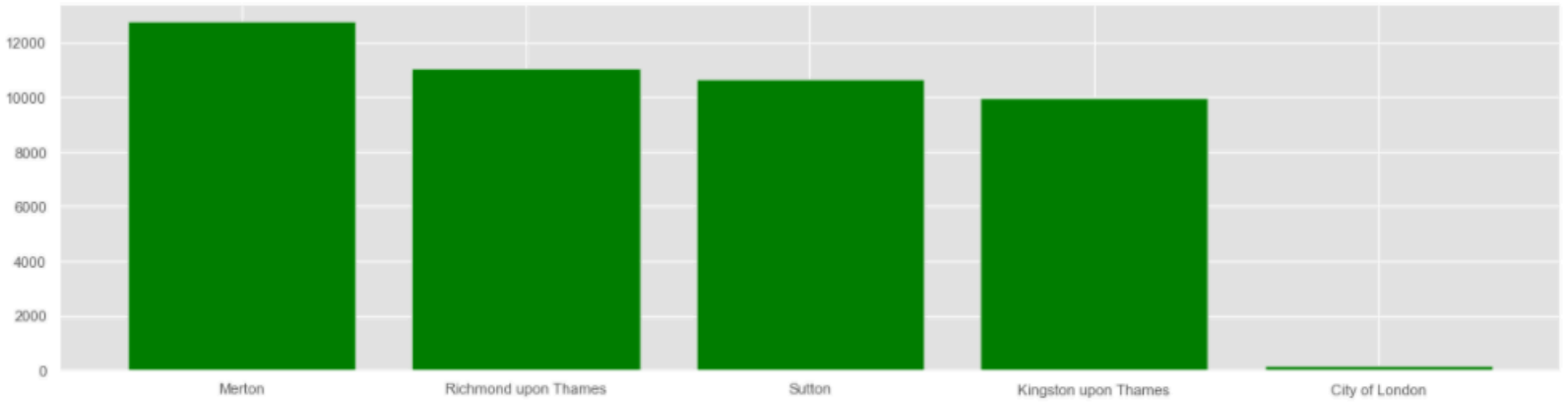
Methodology

- EDA(Top 5)



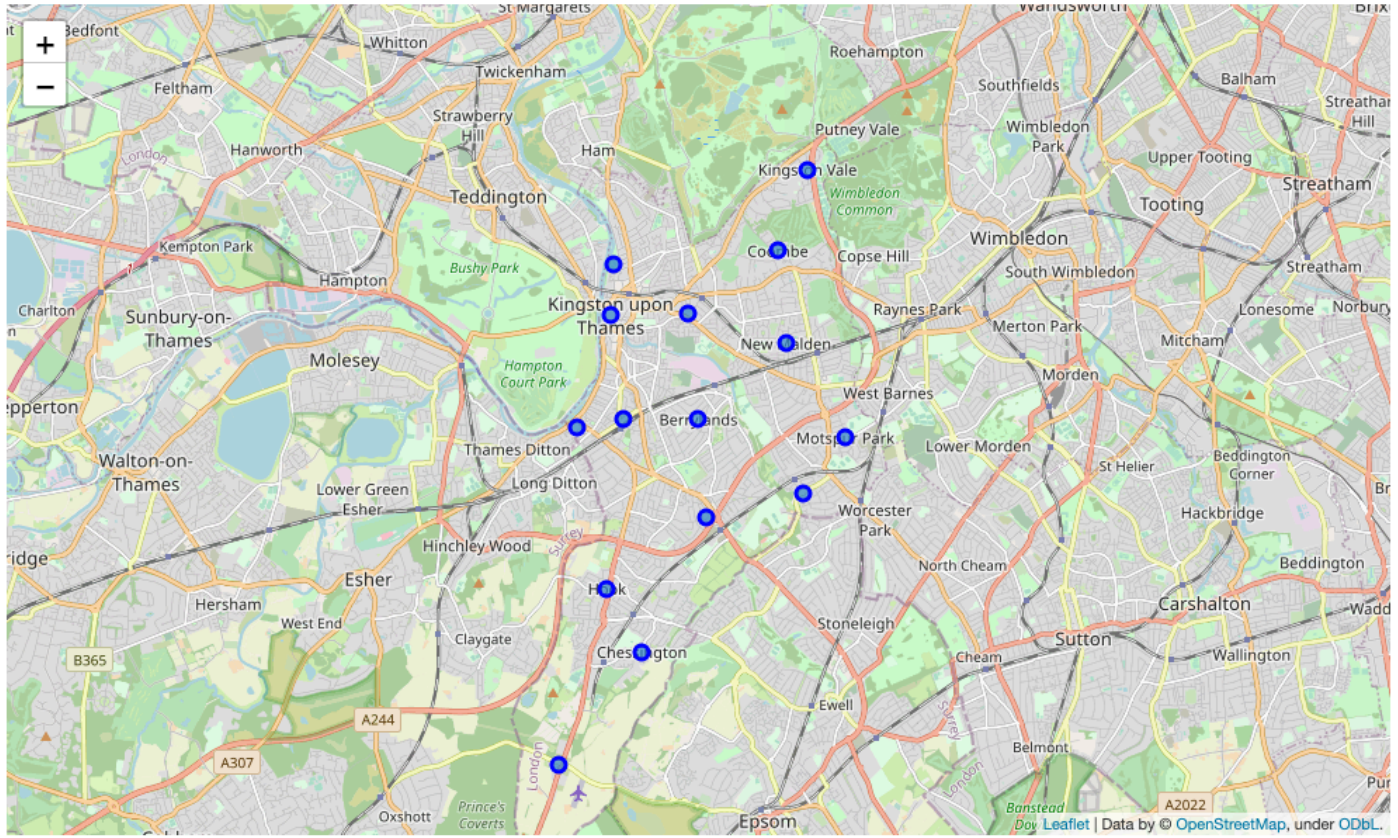
Methodology

- EDA(Low 5)



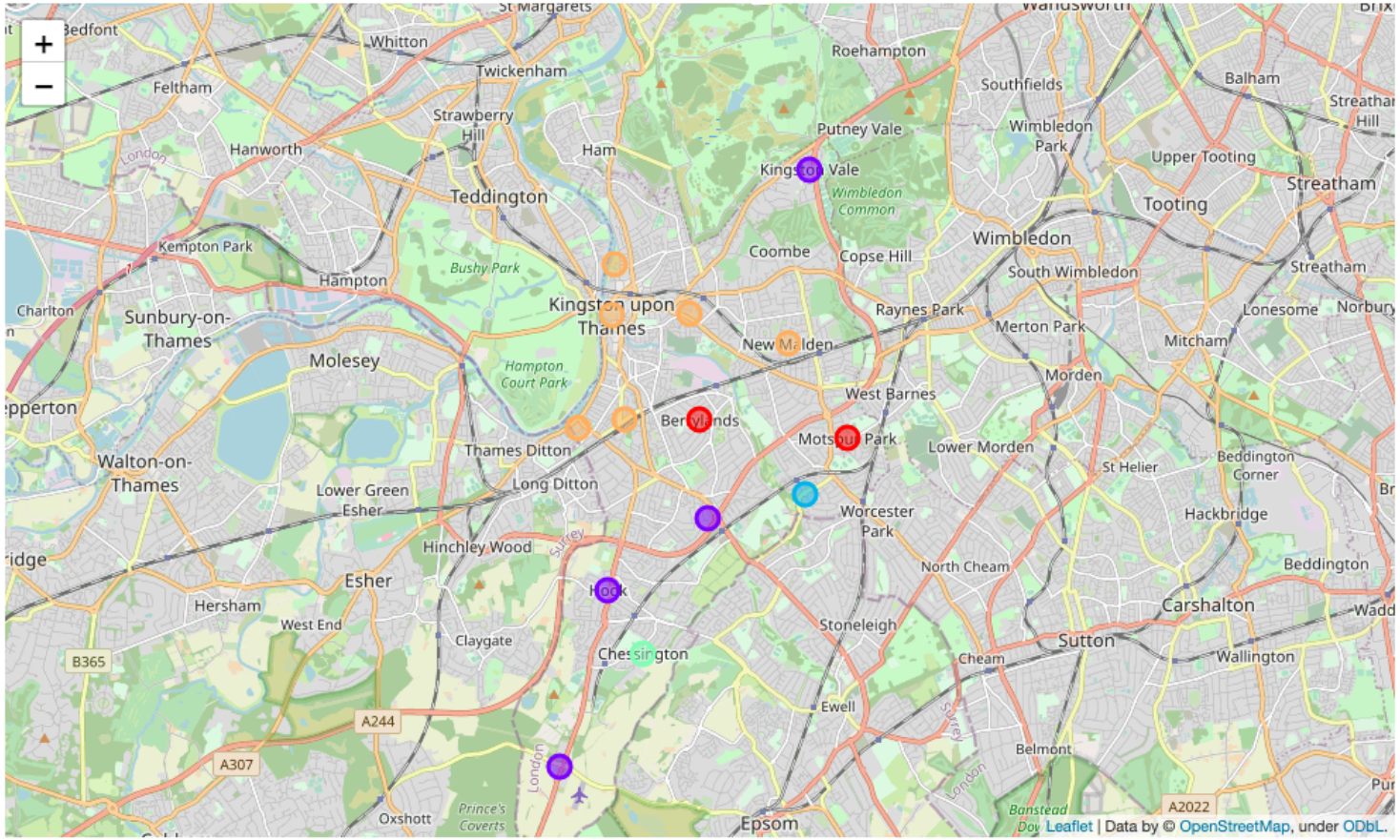
Methodology

- Neighborhoods



Methodology

- Clustering



Result and 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 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. If a person is looking for a neighborhood with stores and restaurants in a close proximity then the neighborhoods in the first cluster is suitable. 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.

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 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 based on safety and a predefined budget.