Introduction: Business Problem

This project aims to select the safest borough in London based on the total crimes, explore the neighbourhoods of that borough to find the 10 most common venues in each neighbourhood and finally cluster the neighbourhoods using k-mean clustering.

This report will be targeted to people who are looking to relocate to London. In order to finalise a neighbourhood to hunt for an apartment, safety is considered as a top concern when moving to a new place. If you don't feel safe in your own home, you're not going to be able to enjoy living there. The crime statistics will provide an insight into this issue.

We will focus on the safest borough and explore its neighbourhoods and the 10 most common venues in each neighbourhood so that the best neighbourhood suited to an individual's needs can be selected.

Data

Based on definition of our problem, factors that will influence our decision are:

➢ The total number of crimes committed in each of the borough during the last year.
➢ The most common venues in each of the neighbourhood in the safest borough selected.

Following data sources will be needed to extract/generate the required information:

➢ Part 1: Pre-processing a real-world data set from Kaggle showing the London Crimes from 2008 to 2016: A dataset consisting of the crime statistics of each borough in London obtained from Kaggle
➢ Part 2: Scraping additional information of the different Boroughs in London from a Wikipedia page.: More information regarding the boroughs of London is scraped using the Beautifulsoup library
➢ Part 3: Creating a new dataset of the Neighbourhoods of the safest borough in London and generating their co-ordinates.: Co-ordinate of neighbourhood will be obtained using Google Maps API geocoding

Part 1: Pre-processing a real-world data set from Kaggle showing the London Crimes from 2008 to 2016

London Crime Data

About this file

o lsoa_code: code for Lower Super Output Area in Greater London.
o borough: Common name for London borough.
o major_category: High level categorization of crime
o minor_category: Low level categorization of crime within major category.
o value: monthly reported count of categorical crime in given borough
o year: Year of reported counts, 2008-2016
o month: Month of reported counts, 1-12

Data set URL: https://www.kaggle.com/jboysen/london-crime