1-Description of Data

For this project, the following data will be used:

- List of London Boroughs with their designations:
  https://en.wikipedia.org/wiki/London_boroughs
- Number of Crimes by London Boroughs:
  https://drive.google.com/file/d/1TTOhuQit4gkenrUizsUn9wloHalRffhE/view

- London Rental Prices by Boroughs:
  https://www.ons.gov.uk/peoplepopulationandcommunity/housing/adhocs/11100privaterentalmarketinlondonjanuary2019todecember2019

- Foursquare API

However, these sources give the data for multiple purposes. This is why they are needed to be cleaned and wrangled accordingly. Some of the columns and row may be needed to be removed, some others may be combined, some nonsense entity may be handled according to need.

First source is a Wikipedia page. Therefore, it is needed to be scrapped by using python library. The second is a csv file which can easily be turned to Pandas dataframe, however it does not have total number of crimes. It has data for different kind of crimes separately from March 2018 to February 2020. So there will be some calculations to generate useable data for our purpose. The third source is an xls file with some of text before the needed table. Also the table is formed in a different structure. It has data for different types of properties separately, instead of an average rental price for each borough. This source has also rental price data recorded between January 2019 to December 2019. So there will be calculation to find average rental prices for each boroughs as well. Final source is Foursquare's API that we used before, during the course. Course's instructions will be followed to obtain venue data for neighbourhoods of selected borough.

London Borough Data Explanation:

Borough: Name of the London Borough.

Designation: Designation of London borough whether an inner London Borough or outher London Borough.

Crime Data Explanation:

Major Text: A general classification that express the type of crime committed.

Minor Text: A little more detailed classification for the crime.

Borough Name: The name of the corresponding London borough.

201xxx: The number of corresponding crime in the year and the month.

Number of Crime: Total number of crime committed in a borough.

London Rental Price Data:

Borough: Name of the London Borough.

Bedroom Category: Type of the property according to number or type of bedroom.

Count of Rents: Number of rented property in a class of corresponding borough and bedroom category.

Mean: The arithmetic average of rental prices in GBP for corresponding property type.

Avg_Rent: Average rental price for a property in a borough regardless of property type.

Foursquare Data Explanation:

<u>Neighbourhood</u>: Name of the neighbourhood in the selected London borough.

<u>Latitude</u>: Latitude value of the location for corresponding neighbourhood

<u>Longitude</u>: Longitude value of the location for corresponding neighbourhood.

<u>Venue</u>: Name of the venue

<u>Venue Latitude</u>: Latitude value of the venue location

<u>Venue Longitude</u>: Longitude value of the venue location

<u>Venue Category</u>: Category of the venue