Optimal location to start a Fast Food Restaurant

City of Choice - **London**

IBM Data Science Professional Certificate - Applied Data Science Capstone

Battle of Neighborhood

20th January 2020

1. Introduction

London is the capital and largest city of England and the United Kingdom. Standing on the River Thames in the south-east of England. It is one of the largest financial centers and has either the fifth or sixth-largest metropolitan area GDP. It is the most-visited city as measured by international arrivals and has the busiest city airport system as measured by passenger traffic. This is a city of opportunities. It is the city of **London**, **we are planning to start a fast-food restaurant**.

1.1. Problem Background

As the city is already crowded with many restaurants and fast-food chains we are going to use the data to find the optimal location to start the restaurant in the Inner London region which is densely populated and with huge floating population.

1.2. Target

We are looking to start a Fastfood restaurant in which we will be targeting tourists who want to grab a quick bite, college students who want to take food , work location where we could provide lunch for the workers, factories where we would get regular customers, etc. also we need to consider the competition and availability of raw materials in the vicinity.

2. Approach

The below section explains the various steps performed in getting the optimal location,

2.1. Data Collection

The data for this project is collected from 3 different data sources,

#	Name of the Source	Link to the Source	Purpose

1	London Data Store	https://data.london.gov.uk/dataset/london-borough-profiles	This provided the details of all the Borough in London along with key parameters like Number of active businesses in that Borough and 2 year business survival in the Borough, Long with that it also provided the Location of the Borough is it part of Inner or Outer London.
2	Wikipedia	https://en.wikipedia.org/wiki/List_of_London_boroughs	This provided some of the key values for the project like Population and coordinates of each of the borough
3	FourSquare	https://api.foursquare.com/v2/venues	Foursquare api is used to collect the places of interest for us both preferable the competition locations are all gathered here

2.2. Tools used

- Python 3.x
    a. Jupyter notebook
    b. numpy
    c. pandas
    d. folium
    e. matplotlib
    f. Scipy
    g. request

2.3. Steps taken

We will the step by step approach

i. Started with consolidating the data from the London Data Store, Wikipedia and add it respective dataframe
ii. Consolidate and merge the data frame with required columns
iii. In the data the Latitude and Longitudes are parsed cleaned
iv. Post that only the Inner London locations are selected
    a. City of London
    b. Camden
    c. Hackney
    d. Hammersmith

e. Haringey
f. Islington
g. Kensington and Chelsea
h. Lambeth
i. Lewisham
j. Newham
k. Southwark
l. Tower Hamlets
m. Wandsworth
n. Westminster

v. Then using the Foursquare API all the attractions are captured based on the Boroughs
vi. As there is a limit in the number results that it fetches we are fetching data based on each type of location.
vii. So a total of 19 different request based on the locations are fetched, Categories considered are -
    a. The details of the categories are listed below
viii. Then the results are grouped based on boroughs listing down the count of each of the category
ix. Then he numbers are categories
    a. *Two-year business survival rates(weighing score 2)* - To be sure this location is good for business
    b. *Population(*weighing score *3)* - To be sure the location is having enough people and that will help business
    c. *Tech Startup(*weighing score *2)* - Gives regular customers
    d. *Supermarket(*weighing score *3)* - Place to shop for daily needs
    e. *Soccer Field(*weighing score*3)* - User to match day crowd
    f. *School(*weighing score *4)* - Gives regular customers
    g. *Pub(*weighing score*2)* - Gives regular customers
    h. *Supermarket(*weighing score *2)* - Place to shop for daily needs
    i. *Office(*weighing score*3)* - Gives regular customers
    j. *Historic* Site*(*weighing score *3)* - Gives tourist customers
    k. *Grocery Store(*weighing score *3)* - Place to shop for daily needs
    l. *Factory(*weighing score*3)* - Gives regular customers
    m. *College Classroom(*weighing score *2)* - Gives regular customers
    n. *Coffee Shop(*weighing score *2)* - Gives regular customers
    o. *Café(*weighing score *2)* - Gives regular customers
    p. *Bus Stop(*weighing score *2)* - Gives regular customers
    q. Fast Food Restaurant(weighing score *-4) negative as its competition*

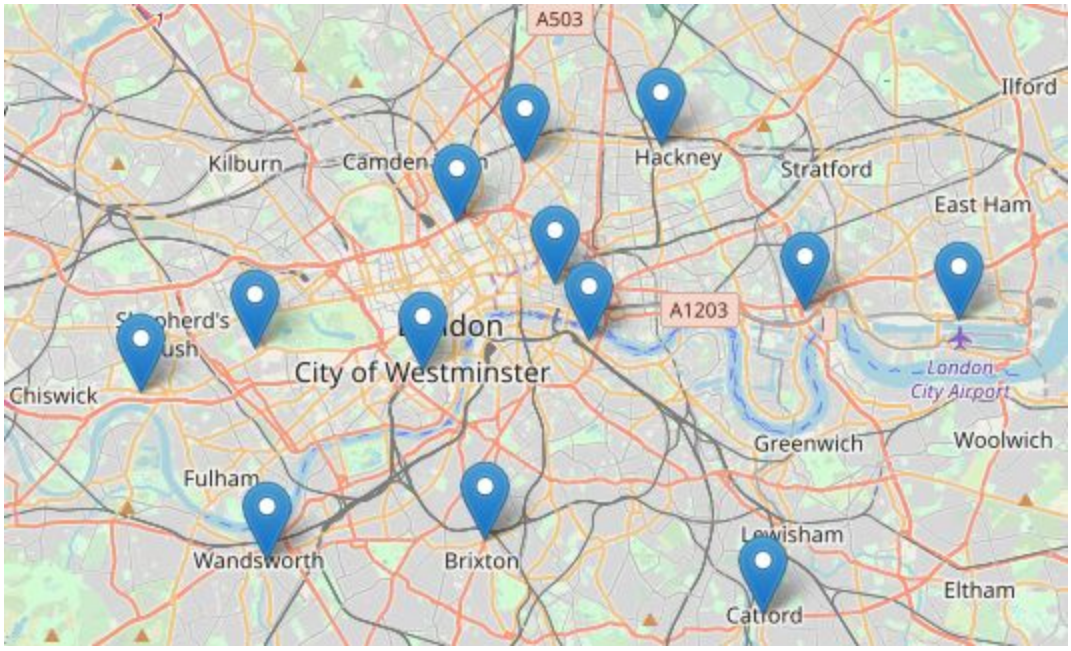
r. Breakfast Spot(weighing score -*3*) *negative as its competition*
s. Sandwich Place(weighing score -*2*)*negative as its competition*

x. Weighted average is calculated for the same
xi. Once this is done it's clearly visible that Westmister stands out as the best option.
xii. So that next step to calculate the KMean was not done as the result was obvious.

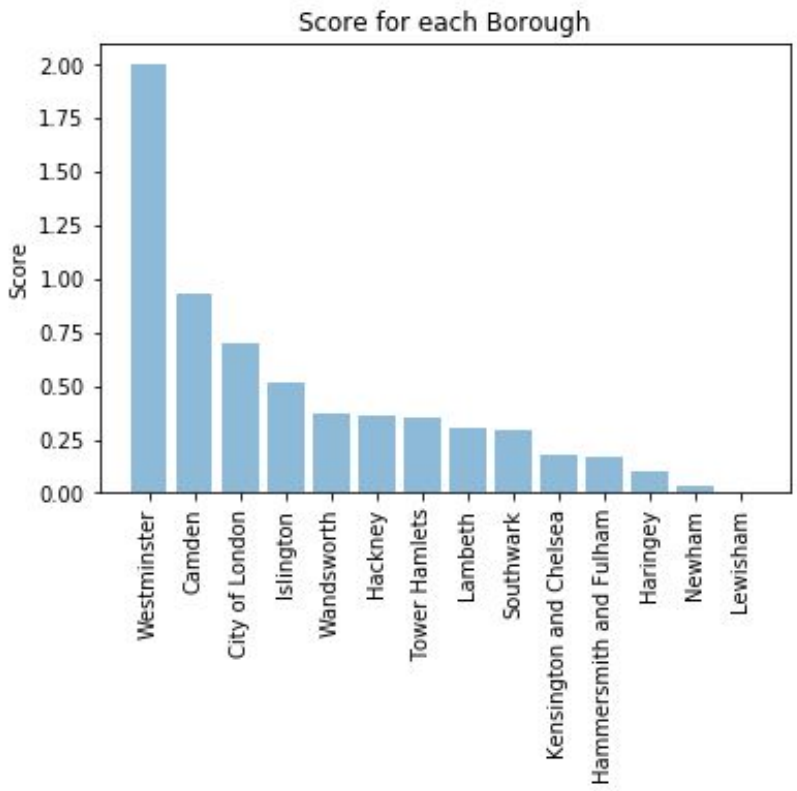
3. Milestone

3.1. Borough narrow down

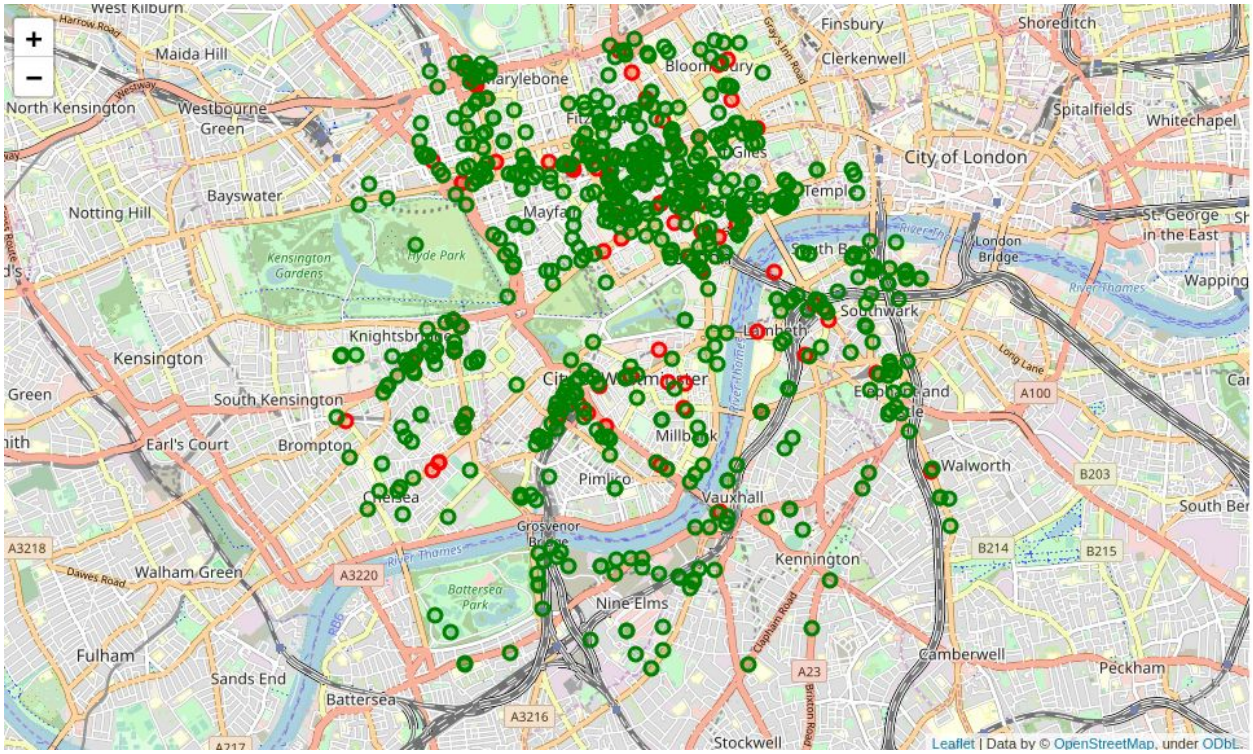
The Boroughs in the Inner London was narrowed as this will high floating population and an opportunity for Fast Food So only 13 Boroughs were selected



Then the parameters were picked based on the options available in the Forusquare. On completing the Weighted average to find the best score for the location based on the Westminster stands out as the best option. Upon checking further in the data it's clear that the business in that are and the Historic attractions in that area is giving a nice benefit for this location



The plot of competition and prefered neighbours are shown in the below plot



4. Assumptions

4.1. Data Set

A very simple dataset was taken for the measure, and even the location set was taken with only 700 per location for each category because of the restriction at the Foursquare end. Also the some of the data that we are currently holding is old population and few other parameters are old.

4.2. Parameters Considered

The data is gathered only 2 selected sources only very few parameters are captured, parameters like rental cost in the Borough, Council restrictions on shops, safety in the area are not taken into account, the best place for Takeaway options was not taken into account.

5. Conclusion and Future enhancements

5.1. Future Improvements

To add many more parameters and to append legal restriction and other Council related benefits from each of the Borough, consider all types of restaurants as competition, get a complete data set, pick a fresh dataset etc, to make a better prediction of the location. Also we could add the distance parameter to each of the location both favorable neighbours and Competitions to even pinpoints a better coordinate.

5.2. Conclusion

From the data that is captured and the parameters that are taken into consideration we conclude that the **Westminster Borough** is the best place to start a **Fast Food restauran**t.

5.3. Reference

1) Wikipedia: https://en.wikipedia.org/wiki/London
2) Data store: https://data.london.gov.uk/dataset