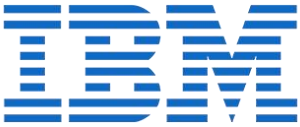
Capstone Project: The Battle of the Neighborhoods (Part 1)

Final Report

By: Abhay Saini; June 21st 2020



Contents

1. The Problem

1.1 Introduction/Background

London is a huge metropolis, probably the one of the oldest cities in continued existence. And since it is the capital of the United Kingdom, one of the first nations which stepped into the Modern Age, it is extensively mapped out and detailed.

There are two locations with the name London, one is the tiny City of London, the heart of the metropolis and the center of the administrative governance. The other is called "Greater London", the bigger, wider area where London's growing development snaked its tentacles to. We would be talking about "Greater London" in this project whenever we speak of London.

Its population is around 8.9 million and area is around 1,572 Square Kilometers. The official language of the city is English but because of England's colonial history, an extremely diverse colonial diaspora has settled inside the city and consequently giving their own flavor to the beautiful city.

1.2 Problem to be Solved

The Problem that we will be going to solve in this Project will be "If someone were to open a new restaurant in London, which area or neighborhood, should he or she choose. The Stakeholders interested in the solution of this problem would be Prospective Entrepreneurs, City Planners and Franchise Owners wanting to Extend a Foreign Franchise into London.

2. Data

2.1 Data Sources

- I took data from primarily one source only. It was London Post Code Data from Doogal.co.uk https://www.doogal.co.uk/london_postcodes.php.
- It consisted of postcodes and other relevant data (Which I have trimmed to only relevant columns, the list of columns which will be used are mentioned below in image 1.0).
- The Analysis will be done at a Ward level, (a collection of postcodes). Since the total number of wards is around ~ 630 and the total area of Greater London is 1,572 sq.-km, average ward size came out to be ~2.2 sq.-km, a perfect size for a neighborhood)
- Secondly, I will use Foursquare Data to get Venue Categories, likes and comments. (The list of columns is mentioned below in image 1.1)

Postcode
Latitude
Longitude
District
Ward
District Code
Ward Code
Region
Index of Multiple Deprivation
Average Income
Distance to station

Postcode data

Venue Category
Likes
Comments
Users Reviews

Foursquare data

Description of the Postcard Data; -

1. Postcode: - Smallest and most basic unit of house identification.
2. Latitude: - Average Latitude of the postcodes of the Ward
3. Longitude: - Average Longitude of the postcodes of the Ward
4. District: - Consists of several Wards
5. Ward: - Collection of postcodes, clubbed together on some administrative commonality
6. Ward Code: - Unique Code of each Ward
7. Region: - East or North or West or North West etc
8. Index of Multiple Deprivation: - Higher the index, more deprived that post code is
9. Average Income: - Income of several households averaged over the postcodes in one ward
10. Distance to Station (Tentative): - Distance to the nearest train station