

# Capstone Project

## The Battle of Neighborhoods (Final Report)

### 1. Introduction

Our stakeholders want us to make a data-driven decision about which is the best neighbor in London to open an Asian Restaurant.

### 2. Data

2.1.1 We will use the data available in London Datastore (free data-sharing portal) in order to decide which borough we will have our Asian restaurant (<https://data.london.gov.uk/> )

- i. This data will help us to find the most populated considering the targeted ethnic group that is interested in our restaurant. Moreover I will consider the population projection till 2030 in order to make sure the business is profitable for the coming 10 years.  
<https://data.london.gov.uk/dataset/ethnic-group-population-projections>
- ii. Another data that will help us is the weekly earning per year in each borough <https://data.london.gov.uk/dataset/earnings-place-residence-borough>

2.1.2 We will use foursquare to check the neighborhoods of the chosen borough. Which locations have the most visited restaurants? Which location is populated by Asian..etc.

### 3. Methodology:

As you will see in the notebook:

3.1.1 To work on the solution, I have used Pandas library to read above data in XLS format and convert them into pandas dataframes. I've done

extensive data cleaning to present the data in suitable format, exploration analysis is done as well (line chart)

3.1.2 I got the geolocations of borough & neighborhood from Wikipedia. 3 libraries have been used to get the coordinates from the URLs (**requests**, **urllib** and **BeautifulSoup**)  
[https://en.wikipedia.org/wiki/List\\_of\\_London\\_boroughs](https://en.wikipedia.org/wiki/List_of_London_boroughs)  
[https://en.wikipedia.org/wiki/List\\_of\\_areas\\_of\\_London](https://en.wikipedia.org/wiki/List_of_areas_of_London)

3.1.3 Folium library is used to visualize the neighborhoods.

3.1.4 I used Foursquare API to explore the neighborhoods in London city. Venues, the venues' categories, working hours ..etc are taken from Foursquare.

3.1.5 I've clustered the neighborhoods using **K-means Clustering** algorithm.

#### **4. Result:**

After examining the created clusters. As a result we recommend that Beckton, Custom House, Maryland, Eastham and Manor Park are the best neighborhoods in Newham borough, to open Asian restaurant. The restaurants has the most common visited venues by the public in these areas. Moreover it's highly populated with Asians.

#### **5. Discussion :**

We could explore all the neighborhoods of the borough and could list the most common venues based on their frequency of occurrence. From these results I can strongly recommend the Beckton, Custom house and few other neighbourhoods as a preferred location for our restaurant , as these areas have the restaurant venue as the most common venue.