**Coursera Capstone Project - Battle of Neighbourhoods**

# PROSPECTS OF NEW RESTAURANT BUSINESS IN LONDON

****

Vidya Rani Puliyangodan Kallaiveedu

**INTRODUCTION**

**BACKGROUND**

Restaurant business is thriving in London due to its diverse culture and London being a hotspot for tourism. Being a Londoner, I have experienced the trouble of finding cuisines of my choice in certain part of the city, leading me to wonder how prosperous the business will be if they carefully plan the location before starting a venture. However rather than going with guess work on the location or cuisines, this project is to apply Machine learning on the data available online to analyse the hotspots for new restaurant locations/cuisines.

**BUSINESS PROBLEM**

London is the [capital](https://en.wikipedia.org/wiki/Capital_city) and largest city of [England](https://en.wikipedia.org/wiki/England) and the [United Kingdom](https://en.wikipedia.org/wiki/United_Kingdom). It is home to many people from different cultures, and the city has a flourishing restaurant scene. The restaurant industry in London is growing exponentially; every street is filled with every variety of restaurant, fast food place, cafes and pub. In such a competitive market, it is of utmost importance that the entrepreneurs wishing to open up a new restaurant business should tactically plan before choosing the location and cuisine.

Below given are the few factors to consider before starting a restaurant in London:

* Population(/and demography) in the particular location
* Are there any Tourist attractions venues nearby
* Competitors in the location, Menu of the competitors

Etc.

**The business problem we are trying to solve** is – If an investor is interested in opening a new eating spot in the capital city London –the model should be capable of providing insight on what would be the best profitable type of restaurant for a given location? Or what would be the best location for a specific type of cuisine?

**Target audience -** is entrepreneurs or investors considering to start a new restaurant or franchise requiring guidance in picking the locality for a particular gastronomy OR cuisine for definite location.

**DATASET AND WORKFLOW**

* London boroughs and locations data from <https://en.wikipedia.org/wiki/List_of_areas_of_London> (Wikipedia). Locations with “Post Town” – London are selected for analysis.
* Use the Geopy and Folium library to get the coordinates of every location and map geospatial data on a London map.
* Foursquare API is used to collect the top 200 restaurants and their categories for each location within a radius 500 meters.

Unsupervised machine learning algorithm K-mean clustering would be applied to form the clusters based on categories of restaurants residing in and around the neighbourhoods. And analyse the top 10 most common restaurants in each cluster.

**APIs:**

**• Foursquare API:** This API has a database of more than 105 million places. This project would use Four-square API as its prime data gathering source. Many organizations are using to geo-tag their photos with detailed info about a destination, while also serving up contextually relevant locations for those who are searching for a place to eat, drink or explore. This API provides the ability to perform location search, location sharing and details about a business. Foursquare users can also use photos, tips and reviews in many productive ways to add value to the results.

Folium- Python visualization library would be used to visualize the neighbourhoods cluster distribution of London over an interactive leaflet map.

**Python packages and Dependencies:**

• Pandas - Library for Data Analysis  
• Geopy – To retrieve Location Data  
• Requests – Library to handle http requests

• Wikipedia – to Scrape the data from Wiki  
• Matplotlib – Python Plotting Module  
• Sklearn – Python machine learning Library  
• Folium – Map rendering Library