

FINDING THE OPTIMAL LOCATION FOR A BUSINESS

Sergey Fotin

May, 2020

Data Presentation

The data that will be used to develop this project is based on two sites:

1. The Foursquare API: This data will be accessed via Python and used to obtain the most common venues per neighbourhood in the city of London. This way, it is possible to have a taste of how the city's venues are distributed, what are the most common places for leisure, and in general, it will provide an idea of what people's likes are.

2. Wikipedia's Ethnic groups in London webpage: This site provides information about ethnicity of population in London which is of great utility to solve this problem. The webpage is scraped using BeautifulSoup4, and the table containing Asian population of London is converted into DataFrame. The data contains information about the immigrant population per borough and per nationality. This data will be analysed in such a way that one could determine the best location of venue/restaurant/other based on people's nationalities. For the sake of simplicity, it will be assumed for this exercise that people's likes vary according to their nationality, and that people from one specific country will be more attracted to place that matches the environment and culture of their own countries, rather than the ones from foreign countries.

You can access the data by clicking this link:

https://en.wikipedia.org/wiki/Ethnic_groups_in_London