# London Neighborhoods Project

# Introduction

|  |
| --- |
| Overview  * In this project, I look at the boroughs in London and see if they can be clustered based on the geolocation data on venues from Foursquare. I leverage the Foursquare location data to explore and compare neighbourhoods in London. * London boroughs are very diverse, yet, one can tell from experience that they serve different functions. For instance, Westminster and surrounding boroughs should be more inclined towards visitors (either for business, political, or cultural purposes) with a prevalence of hotels and dining outlets. The east of London should be more included towards parks, coffee shops, restaurants and other facilities for families and residential life style in general. * The project is aimed to give recommendation for people looking to relocate to London and make decisions based on the availability of venues that are of primary importance to them. For instance, for politicians, an availability of hotels and restaurants and theatres. The target audience for the report are people looking to relocate to London, but are unsure about which borough would be a better choice. The project will produce maps and tables based on which a prospective resident can decide whether they value proximity to theatres and cocktail bars versus proximity to parks and gyms. * Central London have been a centre for economic activity. The functions vary, however, such as theatres (musicals, opera, ballet, dramas) for Covent Garden and government offices for Westminster. Camden, for instance, is characterised by cultural venues and residential neighbourhoods (with shops and gyms). * Other London neighbourhoods are expected to vary. Art museums would be more of interest to visitors, while department stores, gyms, swimming pools are more important for residents with families. |
|  |
| Data  * I have collected data on 32 boroughs in London using the Wikipedia page ('https://en.wikipedia.org/wiki/List\_of\_London\_boroughs‘), which included their names their geographical coordinates. Given that the target tables were online, I used a python scraping library, namely Beautiful Soup. After cleaning and organizing the data, it was saved into a Python data frame. * I also used Foursquare API venues data, collecting data on more than 3000 venues around London. I use this information to cluster boroughs based on the prevalence of the types of venues, such as boroughs where the most popular venue are hotels are likely to belong to the same cluster. The venues data contains names, categories, and geographical coordinates. Venues vary from museums to different types of restaurants. The first dataset contains longitude and latitude coordinates of all boroughs and it is going to be combined with the Foursquare dataset of various venues in those boroughs. |