

Coursera Capstone

IBM Data Science Professional Certificate

The Greater London Borough Analysis

By

Enyinna Echetanna Egejuru

December 2020



The Greater London Borough Analysis

Introduction

Greater London is categorised as a cosmopolitan city being the home to roughly 8.9 million people (Population estimates for the UK, England and Wales, Scotland and Northern Ireland: mid-2019) from diverse backgrounds, London is also one of the top financial capitals in Europe. This city is made up of 33 boroughs each having its own distinct characteristics. Choosing a location within Greater London boroughs to setup shop or to reside can be extremely difficult objective to achieve, however with the help of population data, crime data and location data from Foursquare, we will aid you in making a well-informed decision.

Business Problem

In Greater London, which borough would be a better place to start up a business or live in? By using data science methodology and machine learning techniques like K-Means Clustering the objective of this capstone project is to analyse gathered data and select the best borough in the Greater London which aims to provide an answer to the above question highlighted.

Who would this report interest (Target Audience)

Entrepreneurs set out to either start a business or expand their existing business in the Greater London

Persons planning to move to Greater London and looking for the perfect location that will meet their expectations

Upcoming data scientists that would like to have an insight to some of the most used techniques to obtain data, analyse it and finally, tell a story out of it.

Data Description

For this project, the following data is required

- Population data of Greater London
- List of boroughs in Greater London
- Crime data for Greater London
- Co-ordinate data (Longitude and Latitude) to get the venue data
- Venue data

Data Sources:

1. Population / Location Data

We get population data for Greater London, this data shows the 2019 estimates from <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/populationestimatesforukenglandandwalesScotlandandNorthernIreland/mid2>

[019april2020localauthoritydistrictcodes](#) It contains population information for the United Kingdom broken down by age, this data is filtered to show the London 33 boroughs. For geolocation data we use the geopy package in python to derive the co-ordinate data for the listed 33 boroughs.

2. Venues Data – leveraging on the Foursquare API

Foursquare API provides point of interest information and geolocation based on coordinates provided, we will use this API to get relevant venue location information for Greater London.

3. Crime Data

Criminal activity for each borough will be analysed using Recorded Crime: Geographic Breakdown for Greater London - MPS Borough Level Crime (most recent 24 months) from https://data.london.gov.uk/dataset/recorded_crime_summary. It contains information about the number of crimes at borough levels of London per month, according to crime type from the Metropolitan Police Service.