The Battle of the Neighbourhoods: IBM Coursera Data Science Professional Certificate

IBM Capstone Project

Where in London should an entertainment company open a cinema?

Rebecca Williams

Table of Contents

Introduction

This Capstone Project will analyse data from London to answer the question: As an entertainment company, where in London do you recommend that we open our new cinema? The target audience for this project are companies looking to open entertainment facilities in London who are unsure where exactly their venue should be located. During this project, various techniques will be used to answer the question, including data extraction, data wrangling and cleaning, machine learning algorithms, and data visualisation.

Background

An entertainment company has proposed a plan to open a new cinema in London. The company saw an opportunity for the cinema to be built somewhere in London as it has so far been successful in various other large cities around the world. Although the venue's main function is the cinema, there is also a restaurant and bar should customers wish to go for dinner/drinks before or after watching a film. With a population of 8.9 million, London is a hub for entertainment and the cinema will fit well with the lively scene.

Business Problem

The task the company has put forward is to find out which Neighbourhood in London would be the best place to open the cinema. The cinema will need to be in an area where there are no other cinemas within a reasonable distance to generate more profit for the company. Therefore, the objective of this project is to find the best area in London for the cinema to be located.

Data

The following data will be used for this project

- London location data taken from the Wikipedia page https://en.wikipedia.org/wiki/List_of_areas_of_London. This page includes a list of neighbourhood location, the Borough they are in and other location data. A web scraping method will be used to extract the data from the webpage.

- Latitude and Longitude coordinate data for each of the Neighbourhoods extracted from the Wikipedia page. This locational data is needed for finding nearby venues and plotting maps. This data will be generated using the Python Geocoder package.

- Venue data of cinemas in each London neighbourhood. The Foursquare API will be used to get this data. Foursquare holds a database 105 million places, and provides various different categories of venue data.