New Cinemas, London , UK

Coursera Capstone - Final Project
Bharath

Business Requirement

- Analyse the boroughs and neighbourhoods within the Greater London area and identify borough(s) to open a new cinemas which has good return on investment.

- Look for boroughs with already has established entertainment venues

Approach

- Gather data on boroughs and neighbourhoods with boroughs

- Look for entertainment venues in the neighbourhoods

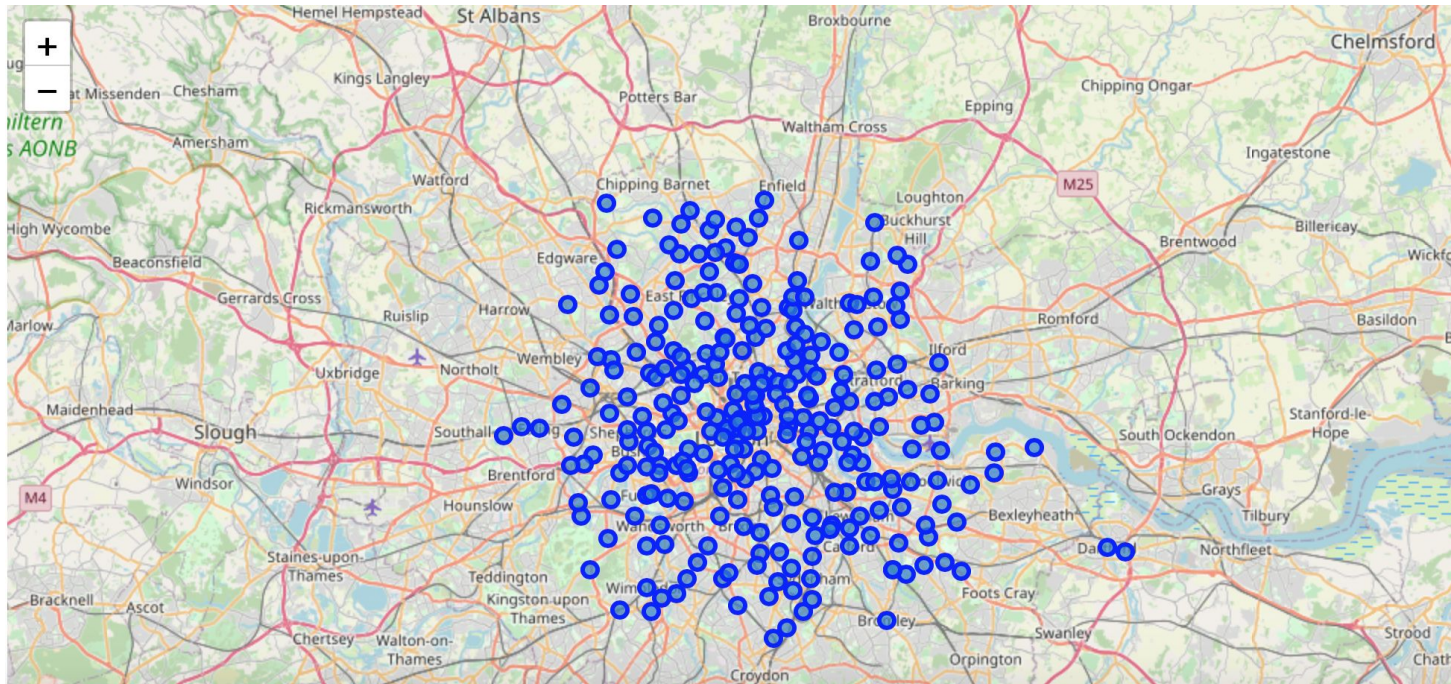- Cluster the venues using unsupervised algorithm, K-Means Clustering

Data

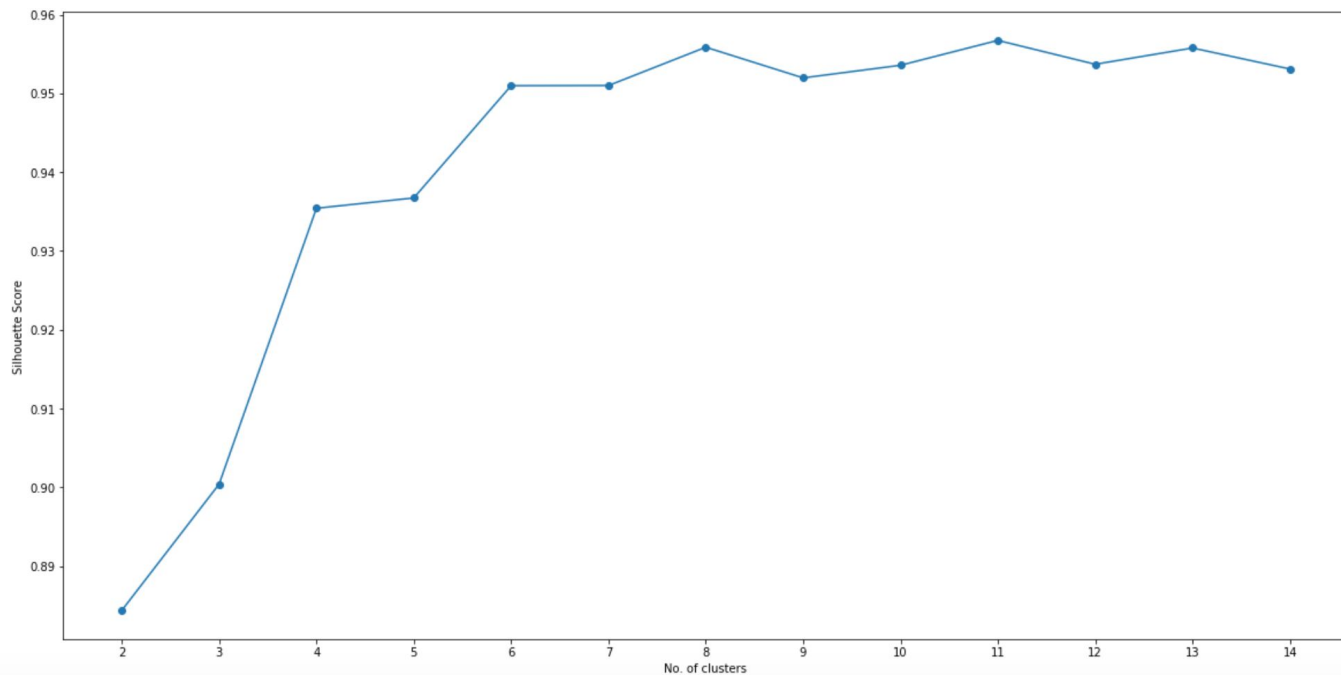| Data	Source
London Neighbourhoods	[Wikipedia](#)
Geo Location	Geopy.geocoders Python Lib
Venues in Neighbourhoods	[Foursquare APIs](#)

Methodology

- Scrape wiki data for Boroughs and Neighbourhoods
- Use Geopy Python library to retrieve geo coordinates for the neighbourhood
- Use Folium library to generate a visual representation of the data
- Determine the optimal K value to be used
- Use K-Means clustering unsupervised algorithm to cluster the data
- Generate a result dataset with cluster labels added.
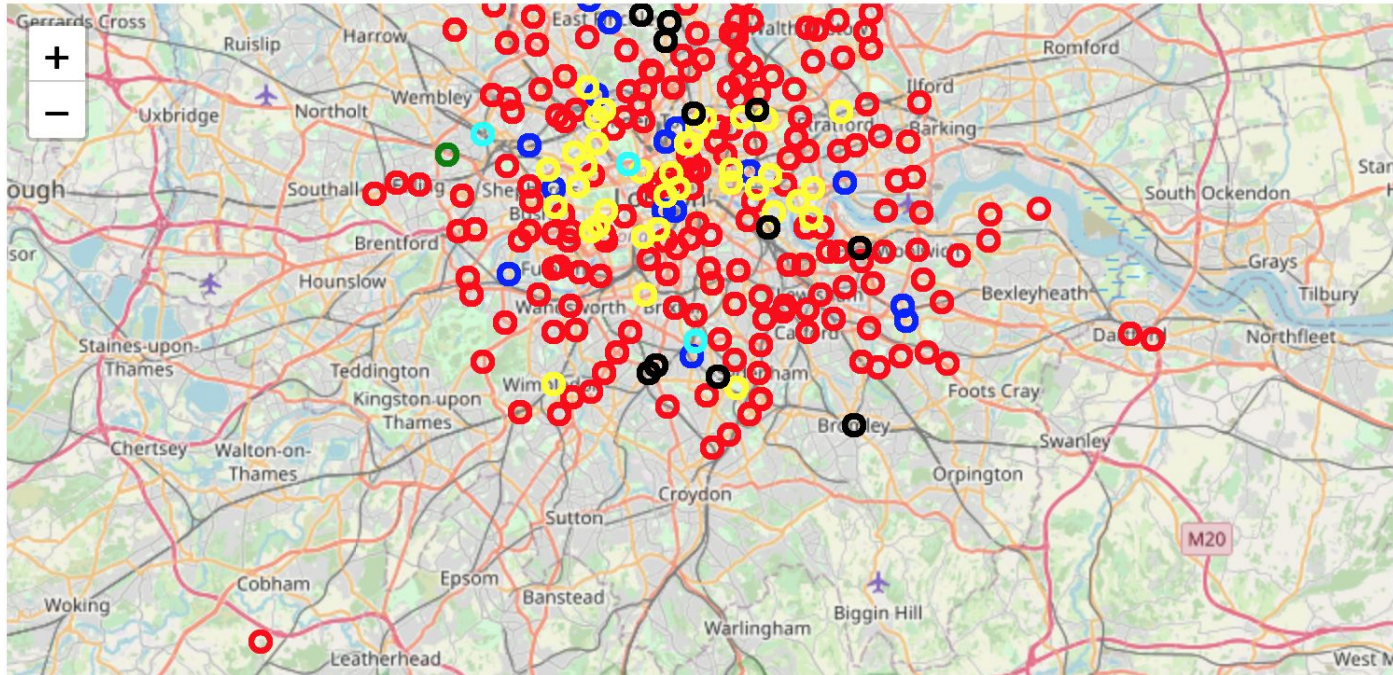
DataSet - Visualised Using Folium

Optimal K - Visualised for Silhouette Score

Clustered Dataset - Visualization

Recommendation

The neighbourhoods in cluster 3 are the most preferred locations to open a Cinemas which are concentrated in Tower Hamlets, Hackey, Westminster, Newham, Islington Boroughs.