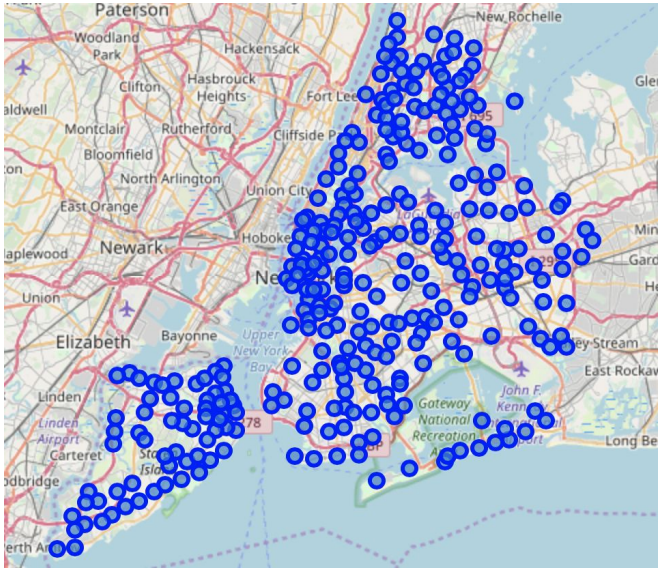
Coursera Capstone Project - Battle of the Neighbourhoods

Presentation

Clustering of neighbourhoods in New York City...

...so as to help new migrants to the city quickly decide their choice of neighbourhoods to live in, based on neighborhood profiles and personal preferences

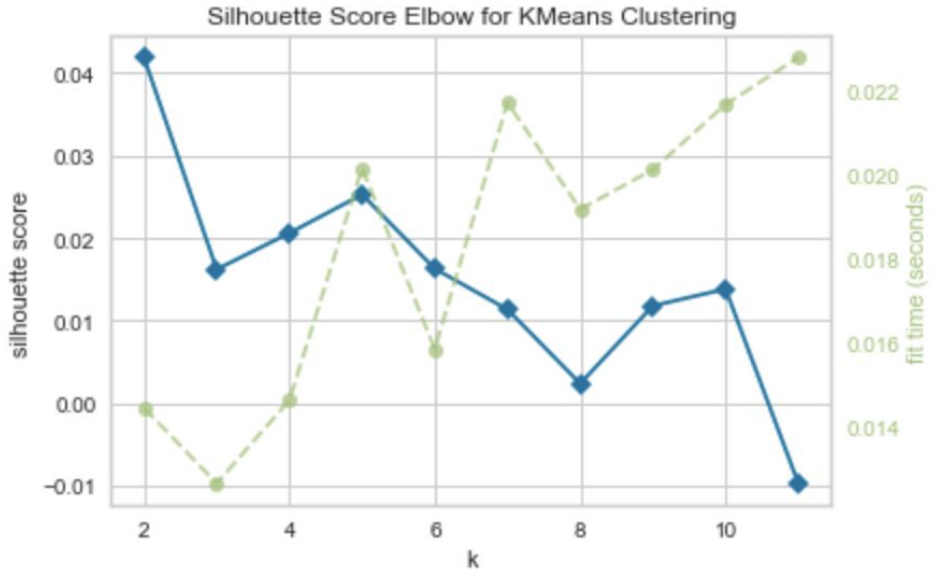


Metrics used for clustering...

1. Average house rent
2. Crime rate (level of safety)
3. Nearby amenities such as restaurants, grocery stores, shopping centres, etc.

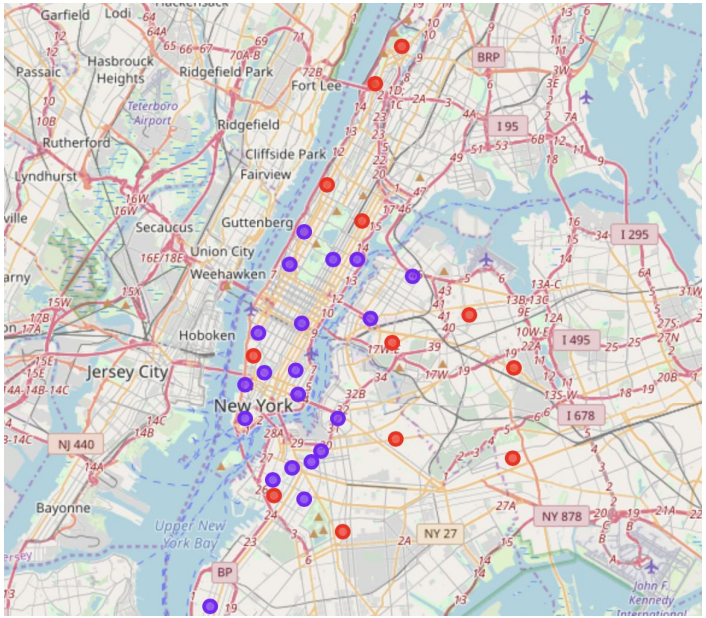
K-means algorithm used for clustering...

...with **number of clusters selected as 2** based on silhouette values calculated



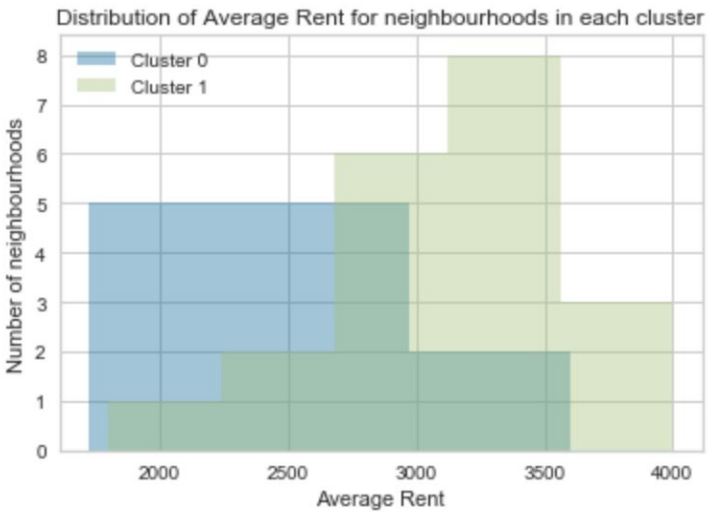
Resulting clusters of neighbourhoods...

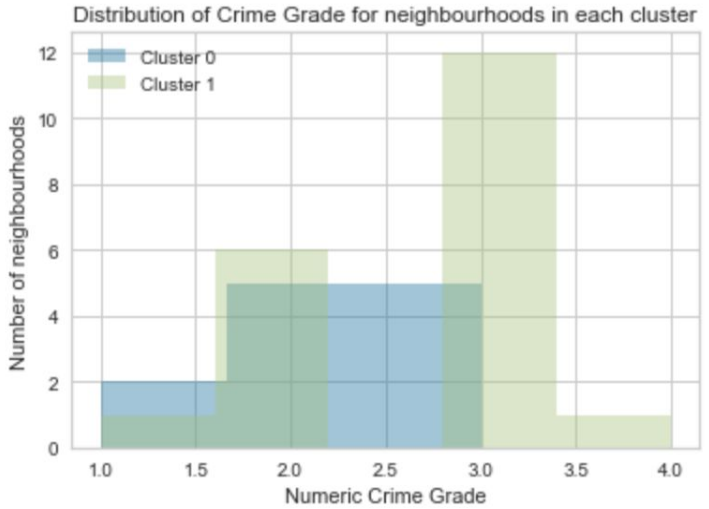
...with **one cluster having 12 neighbourhoods and the other 20 neighbourhoods**



Exploratory analysis of clusters...

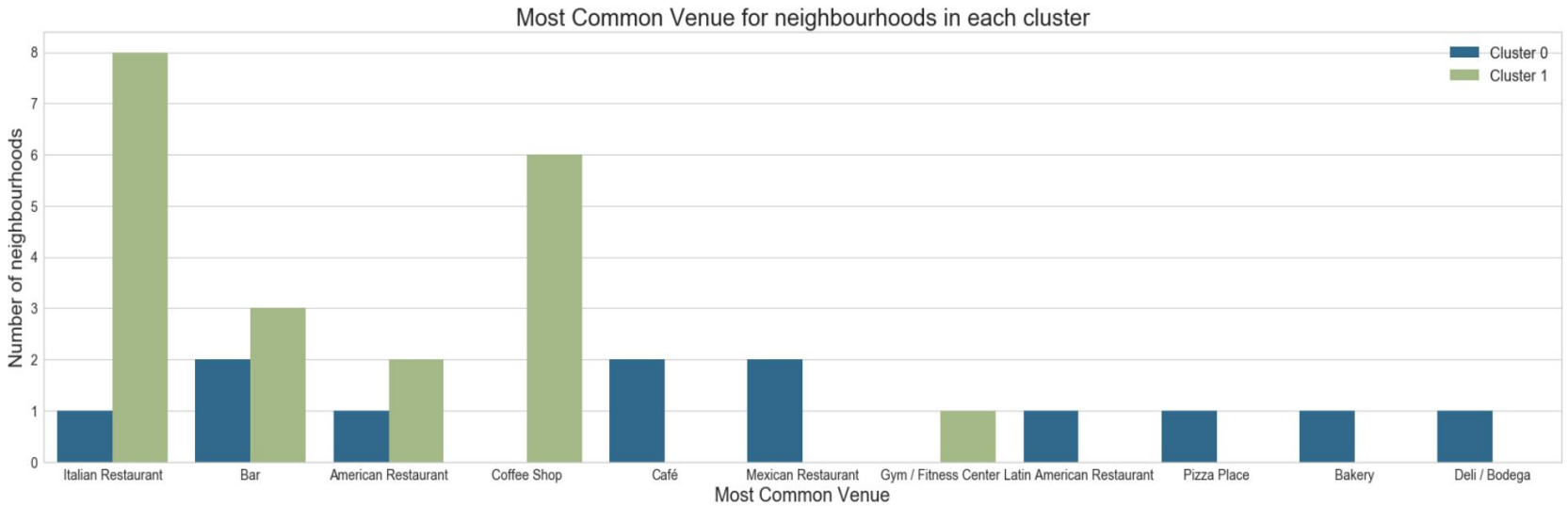
...across average rent, crime rate...





Exploratory analysis of clusters...

...and most common venue categories



Resulting profiles for clusters...

Cluster 0 - Less safe neighbourhoods, with lower rent, and a variety of eateries and cafes

Cluster 1 - More safe neighbourhoods, with higher rent, and hotspots for italian food and coffee shops

In conclusion...

...based on user preferences, the clustering performed enables us to quickly recommend a smaller, more manageable set of neighbourhoods to migrants coming into NYC.