Capstone Project - The Battle of the Neighborhoods

Applied Data Science Capstone by IBM/Coursera

Opening a new Indian Restaurant in Manhattan, New York

By

Divya Singh

Business Problem

- Find an optimal location for an Indian restaurant

- Mainly interested in areas with zero to negligible presence of Indian restaurants

- Objective is to recommend a cluster of neighborhoods to the Stakeholders with none to very little competition

Data Required

- New York City data will be used in this project to explore and analyze the neighborhoods in Manhattan

- Latitude and Longitude of the various neighborhoods in Manhattan

- Venue Data, particularly related to Indian restaurants

Methodology

- Getting New York City data from https://geo.nyu.edu/catalog/nyu_2451_34572

- Use **Foursquare API** to get Venue Data

- Get Latitude and Longitude coordinates using **geocoder**

- Group data by neighborhood and taking the mean of the frequency of occurrence of each venue category

- Filter venue category by Indian Restaurants

- Perform Clustering on the Data by k-means clustering

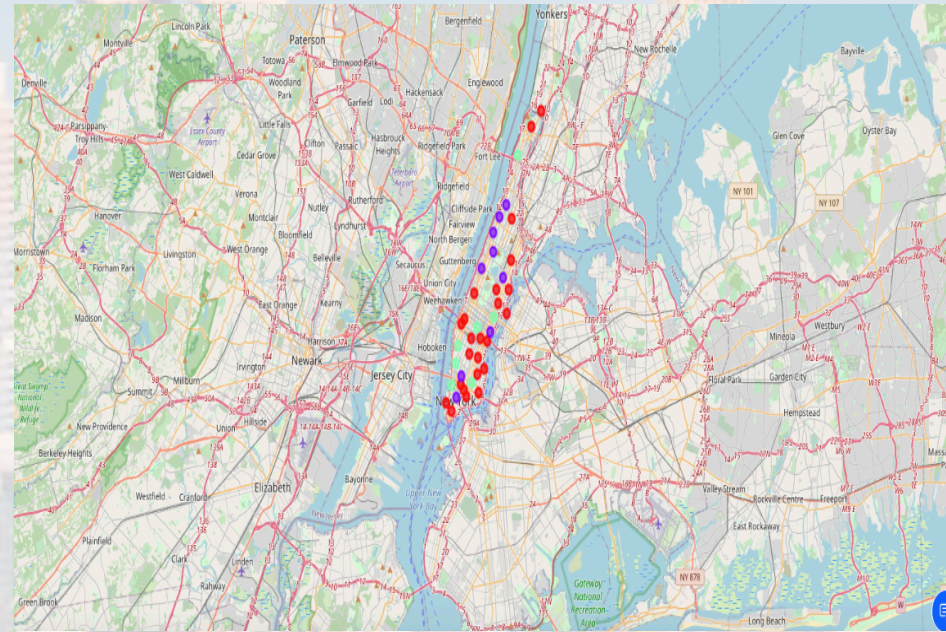- Visualize the clusters in a map using folium

Results

The neighborhoods have been categorized into 3 clusters:

Cluster 0 : Neighborhoods
with lowest/non-existent
concentration of Indian
restaurants.

Cluster 1: Neighborhoods
with highest density of
Indian restaurants.

Cluster 2: Neighborhoods with moderate density of Indian
restaurants.

Discussion

- Cluster 0 could thus be a possible area to consider for opening an Indian Restaurant due to negligible competition.

- This, of course, does not imply that those clusters are actually optimal locations for a new restaurant!

- Purpose of this analysis was to only provide info on areas where the density of Indian restaurants is very low

- It is entirely possible that there is a very good reason for small number of restaurants in any of those areas

- Recommended areas should therefore be considered only as a starting point for a more detailed analysis

Conclusion

- Purpose of this project was to identify neighborhoods in Manhattan, NYC with low number of Indian restaurants

- By calculating restaurant density distribution from Foursquare data, we have identified a cluster of neighborhoods in Manhattan with very low concentration of Indian Restaurants

- Final decision on optimal restaurant location will be made by stakeholders based on specific characteristics of neighborhoods

Thank You