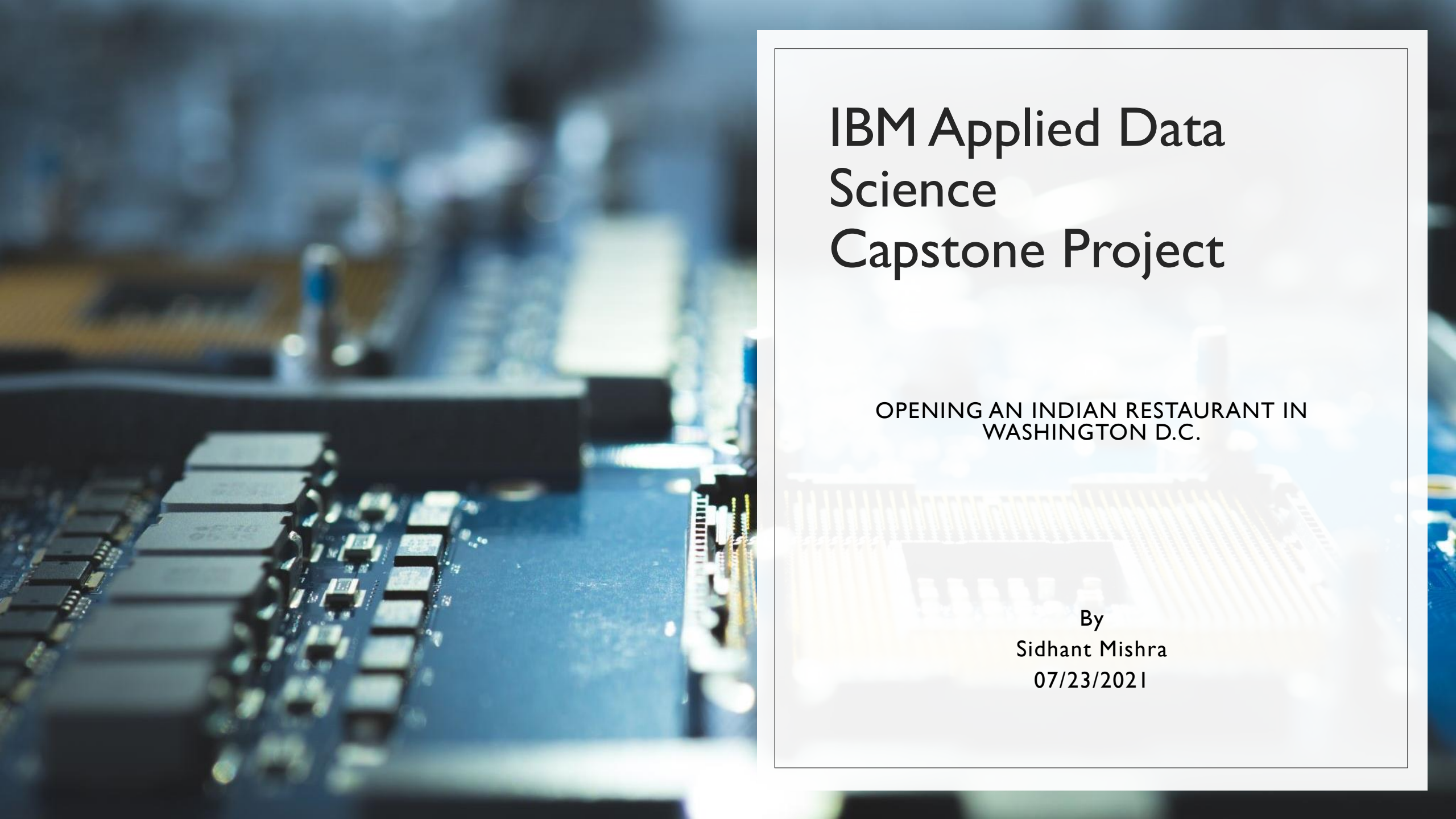


IBM Applied Data Science Capstone Project

OPENING AN INDIAN RESTAURANT IN WASHINGTON D.C.

By
Sidhant Mishra
07/23/2021

Business Problem

- A client having little to no experience in the restaurant and hospitality business seeks to open a restaurant specializing in Indian cuisine in Washington D.C. area. Being from the DC area himself, he kind of understands the demand which Indian food has and certainly believes that given the right strategic location, he could make it a successful one.

- Taking into account the price level at which the restaurant will operate, the intent is to find an optimal location in the DC neighborhood, where there is significant demand for other cuisines, and which already has a good footfall that can be capitalized.

Data Sources

To perform this analysis, the following set of data would be required

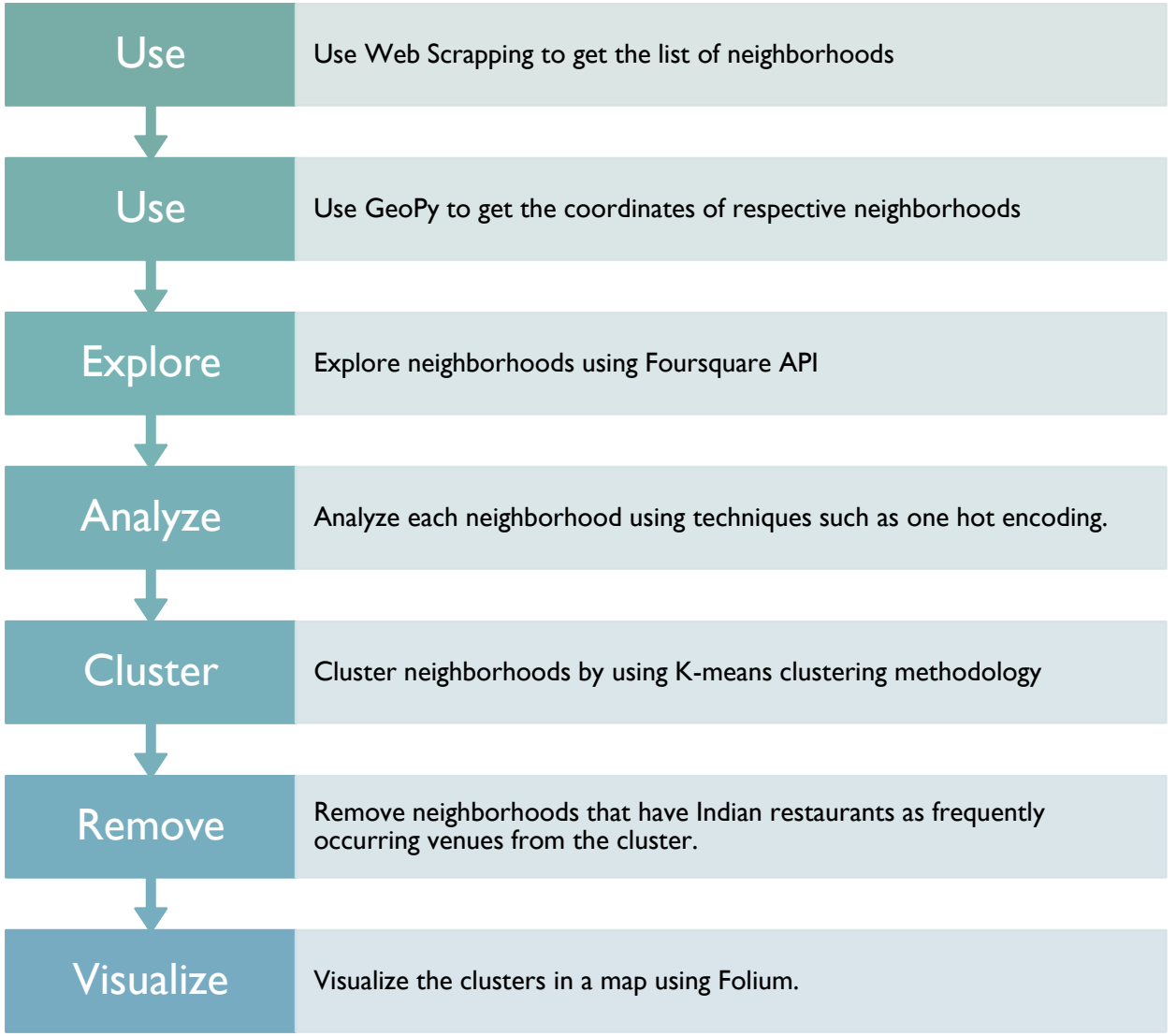
- List of Washington D.C. neighborhoods - The list of neighborhoods can be scraped from here
- Geo coordinates of all the neighborhoods - Geographical coordinates for each neighborhood can be obtained using the geocoder tool in the notebook.
- Top venues in each of the neighborhoods - Data pertaining to top venues would be retrieved using Foursquare API

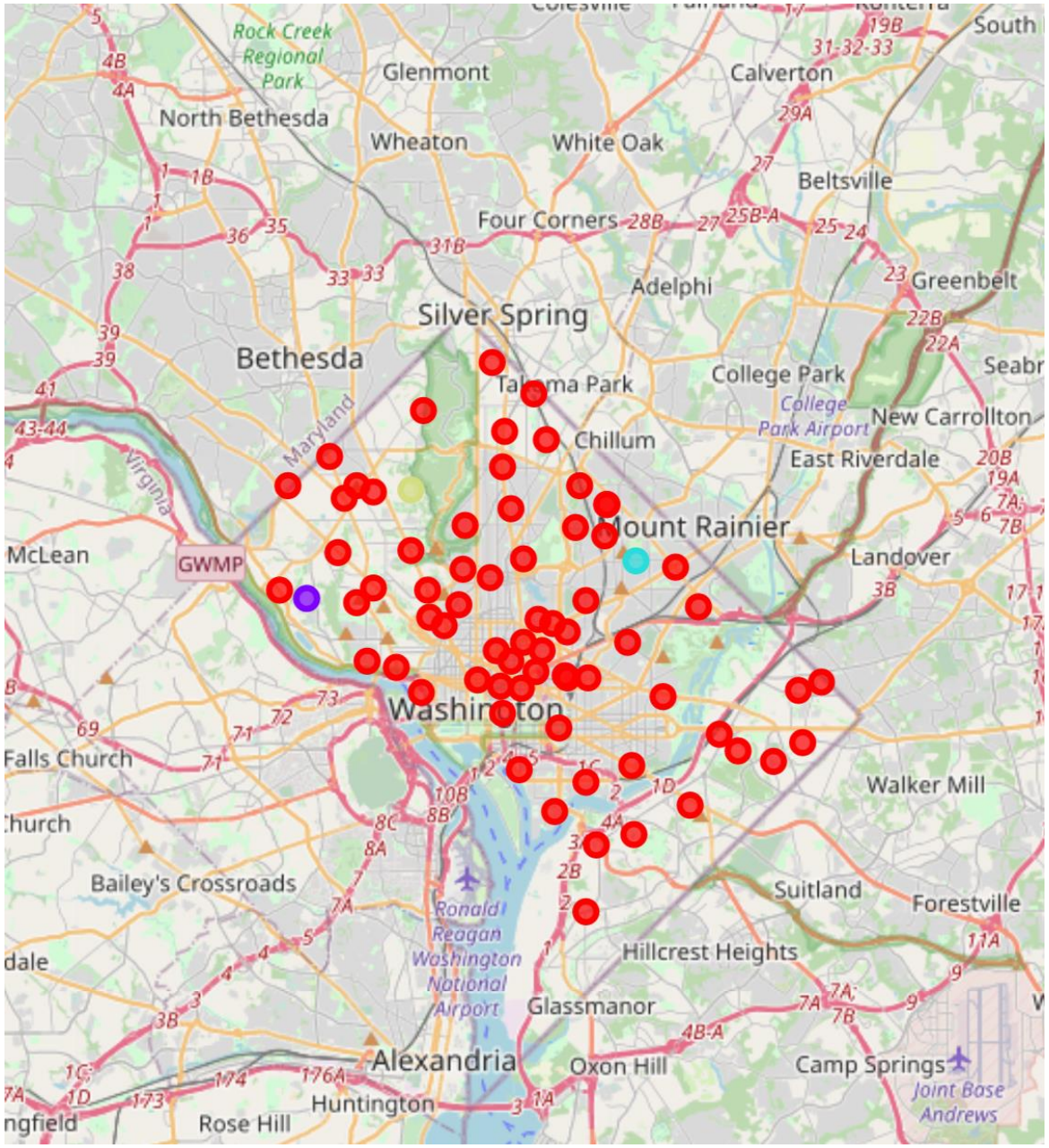
Business Logic

Since we are looking for the most optimal neighborhoods as our probable locations for the restaurant, it is important that we place some parameters on the basis of which we are going to assess neighborhoods/clusters.

- Using the Foursquare API's explore function we would be able to return the neighborhoods that have frequently occurring Indian restaurants. The higher the frequency, the more the competition. The assumption of our analysis is that the barrier of entry to establish a new restaurant in a competitive market is high as existing Indian restaurants may have the competitive advantage of brand loyalty. Therefore, we would not be exploring such neighborhoods for this particular analysis.
- We would be prioritizing neighborhoods that already have presence of cafes, restaurants specializing in other cuisines, bars, coffee shops etc. Such neighborhoods would already have a guaranteed footfall and we can easily capitalize on it.

Methodology





Results

Understanding the clusters

Cluster 1 (Marked in Red)

Our first cluster contains cafes, coffee shops, restaurants of different cuisines, pubs etc. This is the most dominant cluster and contains most of the neighborhoods of Washington D.C. and would be ideal to open a restaurant.

Cluster 2 (Marked in Yellow)

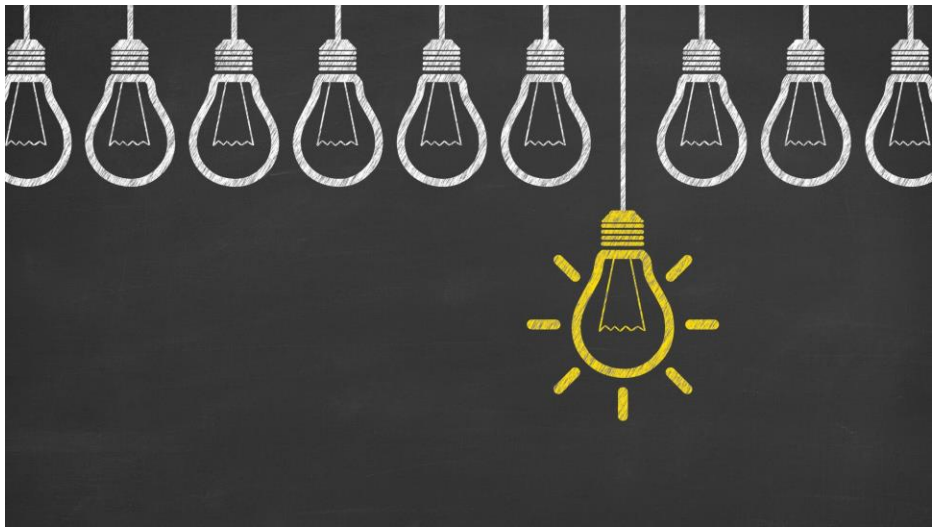
This cluster has just one neighborhood that has mostly museums and exhibition centers and hence wouldn't be ideal for a restaurant business.

Cluster 3 (Marked in Blue)

This cluster has just one neighborhood that has houses and exhibition centers and appears to be a residential area, hence wouldn't be ideal for a restaurant business.

Cluster 4 (Marked in Purple)

This cluster has just one neighborhood that has art exhibition centers and event spaces, hence wouldn't be ideal for a restaurant business.



Discussion

By analyzing each of the clusters, it is evident that cluster 1 appears to be the promising one for the new restaurant to be opened. By looking at the venue category type, it's clear that the cluster contains all the commercial neighborhoods that are busy areas where people come for work, tourism etc. Hence considering the heavy footfall and presence of other businesses, the neighborhoods in this cluster appear to be the best ones for the restaurant to be opened.

Conclusion

- In this project, we have gone through the process of identifying the business problem, specifying the data required, extracting and preparing the data, analyzing and clustering the data into 4 different clusters using various machine learning techniques based on the venue category types and finally visualizing them using Folium map.

- We were able to provide recommendations to our client on the ideal neighborhoods for opening an Indian restaurant. The findings of this project would help potential restaurateurs to capitalize on opportunities present in high potential locations while avoiding competition at the same time.

THANK YOU