IBM APPLIED DATA SCIENCE CAPSTONE

OPENING A NEW INDIAN RESTAURANT IN DELHI, INDIA

BY : PULKIT BHAMBRI

INTRODUCTION

- **Business problem:** This project aims to segment the neighbourhoods of Delhi on the basis of presence of Indian restaurants in them so that the perfect neighbourhoods for the opening of a new Indian restaurant can be determined.

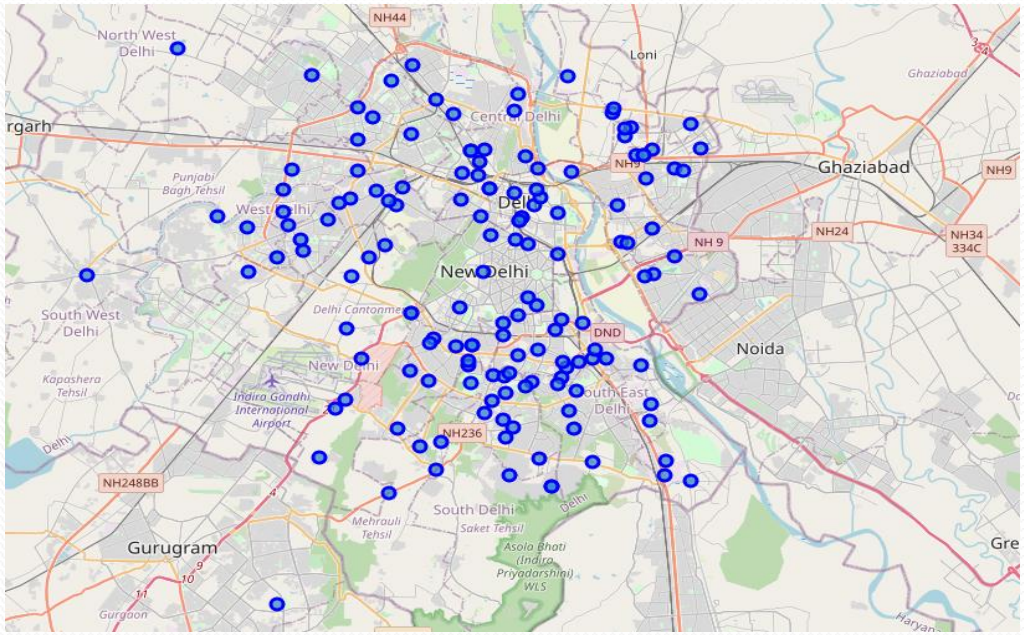
- **Why? :** Location of the restaurant is one of the most important factors that will determine if the restaurant is a success or not.

- **Target Audience:** Any business owner or an individual looking to open a new Indian restaurant in Delhi would be very interested in knowing the best location for his/her restaurant. Existing restaurant owners which aim to make more profits can use this model for relocation of their restaurants as well.

DATA

- Dataset containing the neighbourhoods of Delhi along with their latitude and longitude coordinates is obtained from Kaggle.com

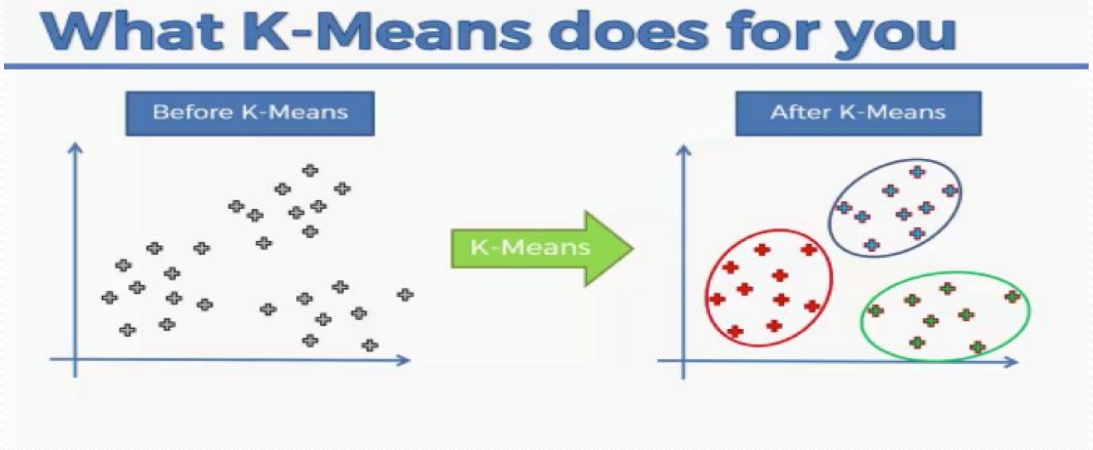
- Venues in a specific radius of the neighbourhoods are obtained by leveraging the Foursquare API



Plot of neighbourhoods of Delhi

METHODOLOGY

- Downloading the neighbourhood dataset from Kaggle.com
- Converting the dataset into a pandas dataframe and cleaning it
- Using Foursquare API to get the top 100 venues in a radius of 1000m from the neighbourhoods.
- Plotting the neighbourhoods on a map of Delhi using the folium package.
- Segregating the Indian restaurants from the venue category
- Applying K- means clustering on the data to segment the neighbourhoods into 4 clusters on the basis of occurrence of Indian restaurants in them.

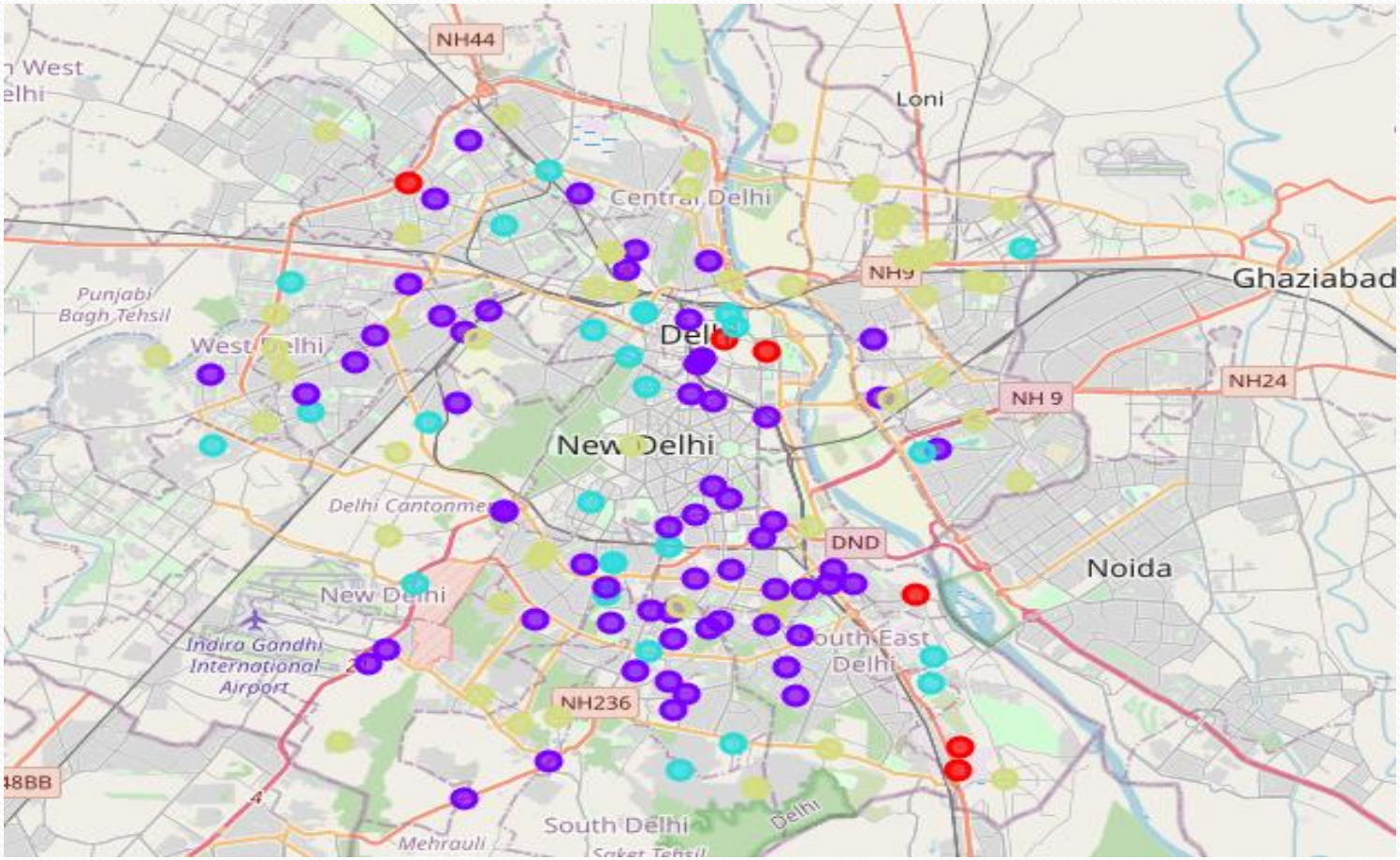


RESULTS

- The neighbourhoods were divided into the following 4clusters:

1. **Cluster 0:** Neighbourhoods with highest number of Indian Restaurants
2. **Cluster 1**: Neighbourhoods with low number of Indian Restaurants
3. **Cluster 2:** Neighbourhoods with moderate concentration of Indian Restaurants
4. **Cluster 3:** Neighbourhoods with least number of Indian Restaurants

CLUSTERS:



Cluster 0: RED, Cluster 1: PURPLE, Cluster 2: Blue, Cluster 3: YELLOW

DISCUSSION

- Based on the results we can see that neighbourhoods in cluster 3 have the least amount of Indian restaurants so these neighbourhoods seem ideal for opening a new restaurant as it would not face competition from other restaurants of the same cuisine.

- Cluster 1 has a low number of Indian restaurants too and can be considered as well. Choice of clusters can depend on the availability of space as well as the type of population in those areas

- Business owners are advised to avoid the neighbourhoods of cluster 0 as they have the highest concentration of Indian restaurants.

- The same analysis can be applied to any category of venues obtained from the Foursquare API be it a new Arcade or a Shopping Mall.

CONCLUSION

- To answer the business question that was raised in the introduction section, the answer proposed by this project is: The neighbourhoods in cluster 3 are the most preferred locations to open a new Indian restaurant.

- The findings of this project will help the relevant stakeholders to capitalize on the opportunities on high potential locations while avoiding overcrowded areas in their decisions to open a new Indian restaurant.

REFERENCES

- Neighbourhoods of Delhi dataset from Kaggle.com:

https://www.kaggle.com/shaswatd673/delhi-neighborhood-data/data?select=delhi_dataSet.csv

- Foursquare API:

https://developer.foursquare.com/docs/places-api/