Coursera – Capstone Project

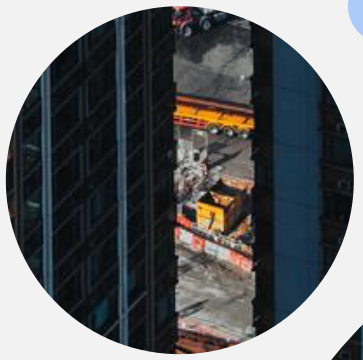
The Battle of Neighborhoods

December 28, 2020

Sunil Kunchoor Basavaraju

Outline

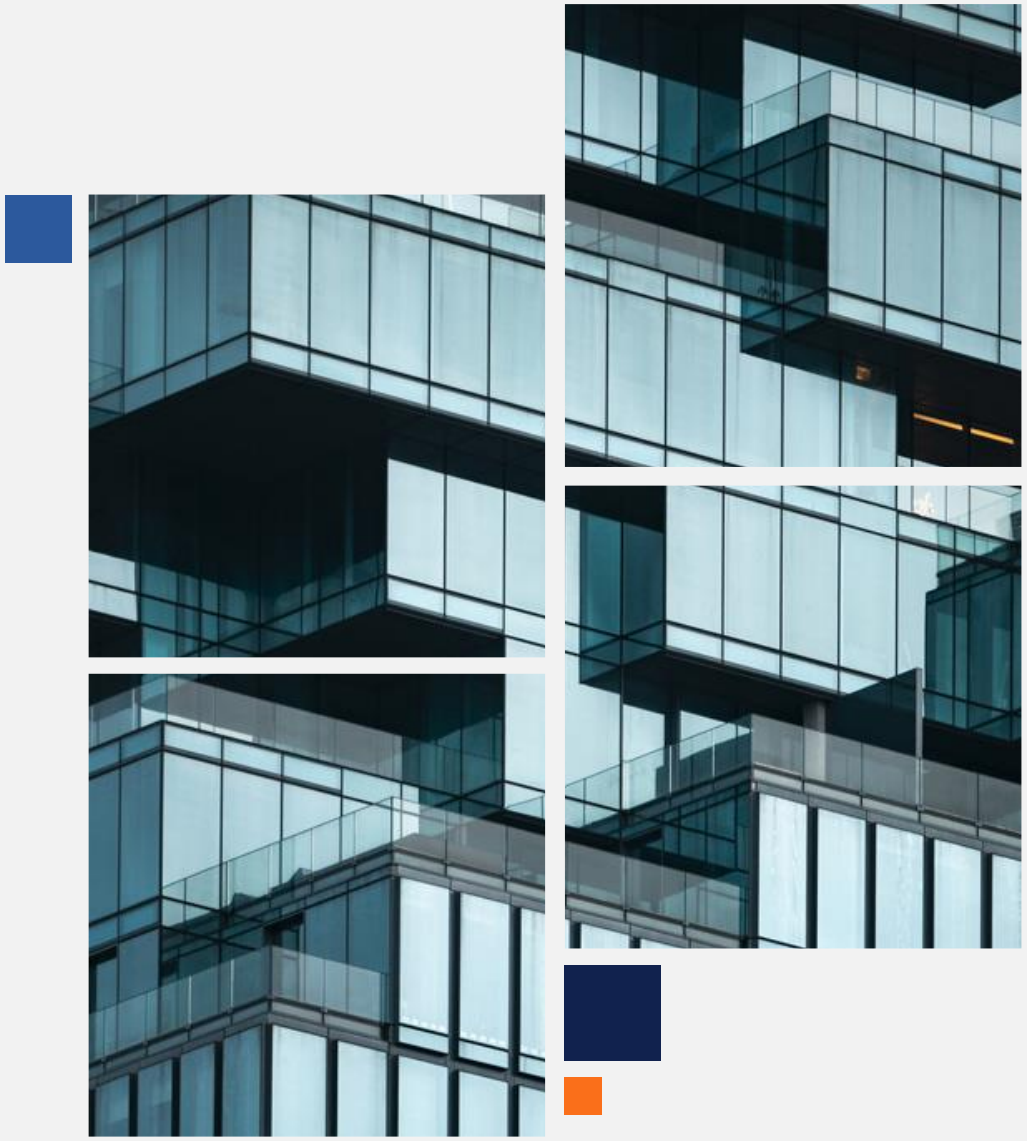
- Introduction
- Business Problem
- Data
- Methodology
- Analysis
- Results and Discussion
- Conclusion
- Appendix



Introduction

In this capstone project, I am utilizing the skills and tools I learnt from Coursera courses.

I have selected the **New York** city for the project. I am helping the stakeholders to narrow the location for their new business.

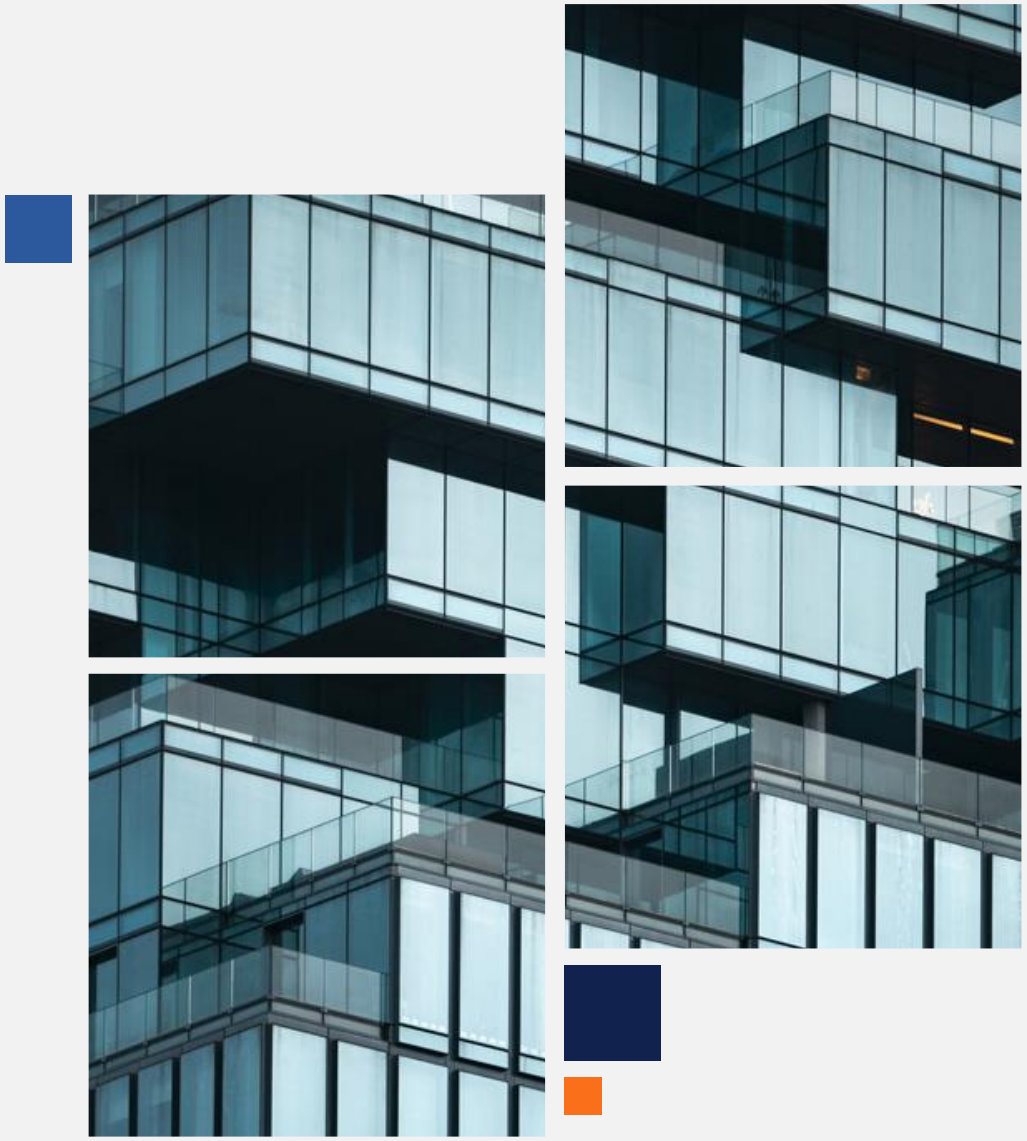


Business Problem

In this project, we are assisting a big Indian restaurant chain to open a new restaurant on a foreign land. Currently, our stakeholders own around 200 restaurants in India and they want to expand their business by opening a new and the first restaurant in the New York City. We have to find a solution for stakeholders to open Indian restaurant chain in the city New York, USA.

Since there are lots of restaurants in New York city, we will find the locations that are not crowded with Indian restaurants and we are also interested in areas with less Indian restaurants.

The stakeholders are not only interested in the location to open new chains; they are also interested in the place where they can make good profit. So, we are helping them to find a location where the areas are crowded with other categories like Art & Entertainment, College & Universities, and Profession offices.



Data

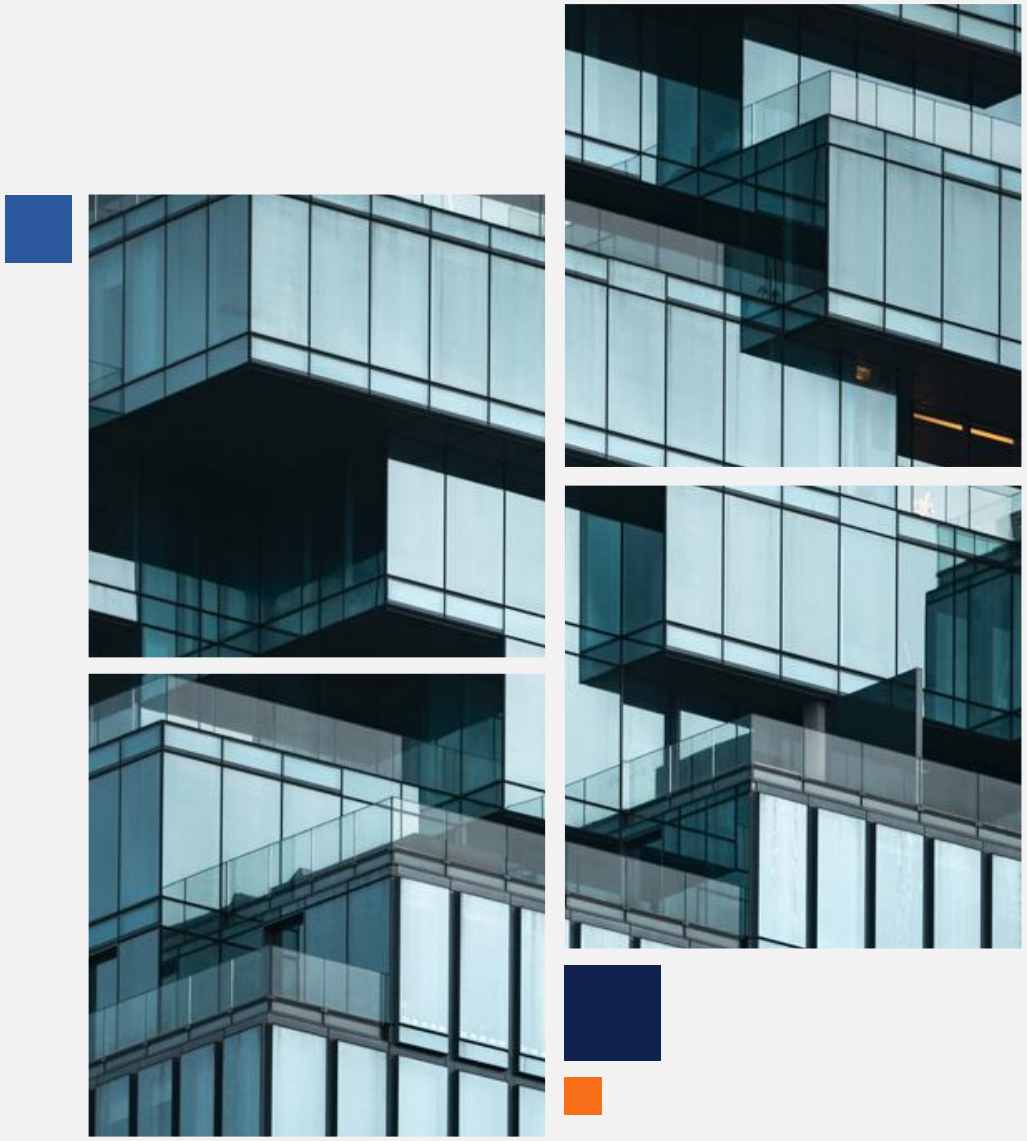
There are five boroughs in New York. We are looking at the number of Indian restaurants in the all five boroughs. Based on the count and crowd of the restaurants, we can make decision on the location to open new restaurant.

Following factors will influence our decision:

- Number of Indian restaurants in the boroughs
- Density of area with other category venues (Art & Entertainment, College & Universities, and Profession offices)

Following data sources will be needed to extract/generate the required information:

- Using geopy python library, I will collect the latitude and longitude for all five boroughs.
- Restaurants, restaurants type, Postal code, and address will be obtained using Foursquare API
- Other business data will also be obtained from Foursquare API. This data will help us to understand the density of the area.
- From Department of Health, NY, borough and neighborhood data will be obtained.



Methodology

Let’s dive in

Methodology

In this project, our aim is to detect areas that have low Indian restaurant density and good place for the start of new restaurant that can bring good profit for the stakeholders.

In the first step, we will collect the all Indian restaurants for each borough using foursquare API.

In the second step, we will combine all the boroughs data into a single DataFrame and displayed them on the map. In this step, we got the picture of Indian restaurants density in the New York city.

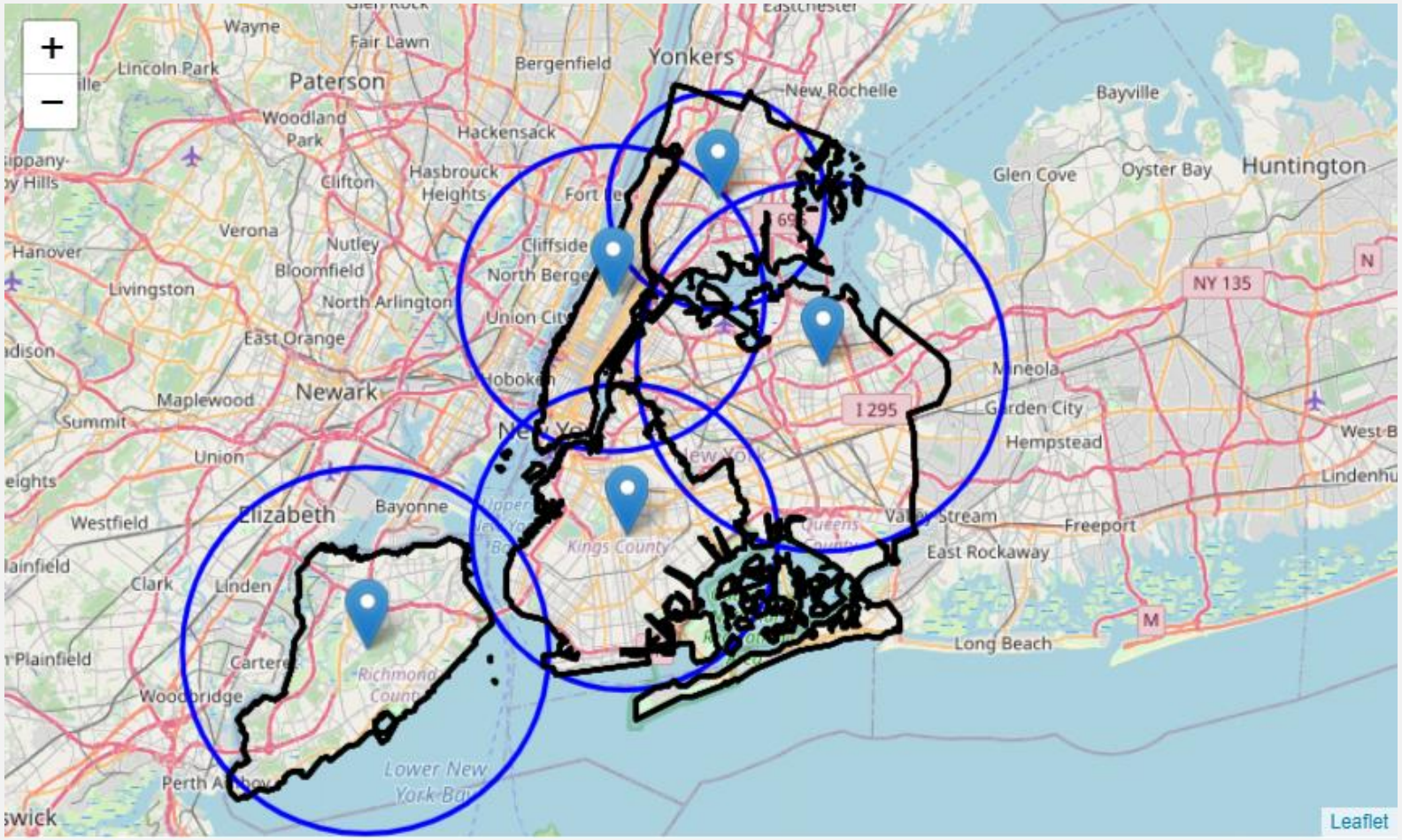
In the third step, we will focus on the creating clusters using unsupervised learning algorithm (**k-means**). The clusters will give an idea about density of Indian restaurants in all areas. We will avoid those areas and we will focus on the area with less dense with Indian restaurants.

In the last step, we will look for other business in the area with less dense Indian restaurants. Because opening a new business at unknown location is not a good business strategy. Before we jump to conclusion on the place to open for business, we will take a look at the venues of categories like **Arts & Entertainment**, **College & University**, and **Professional & Other Places**. We can suggest our stakeholder to open the business at these places.

Analysis

Let’s dive in

Analysis



Using the **geopy** and **Folium** library, we have drawn circles are created to understand the search radius for the Indian restaurants.

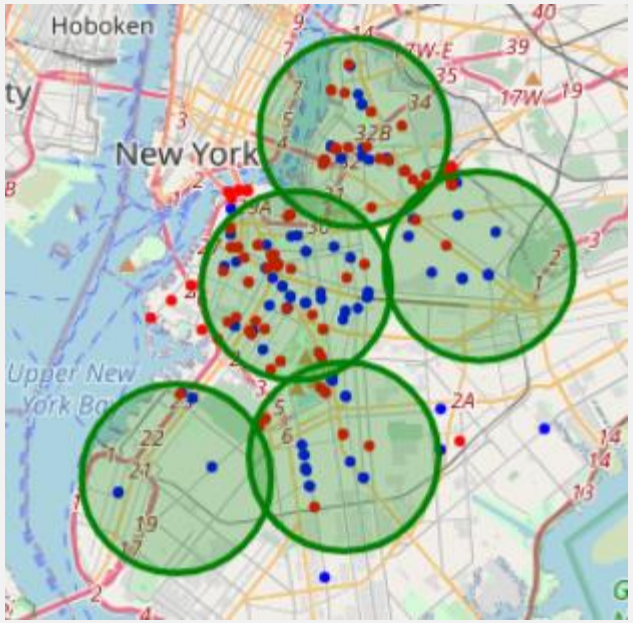
Looks like we are covering almost all the area of boroughs. Though it looks like few are overlapping, we will remove those duplicate data.

Analysis

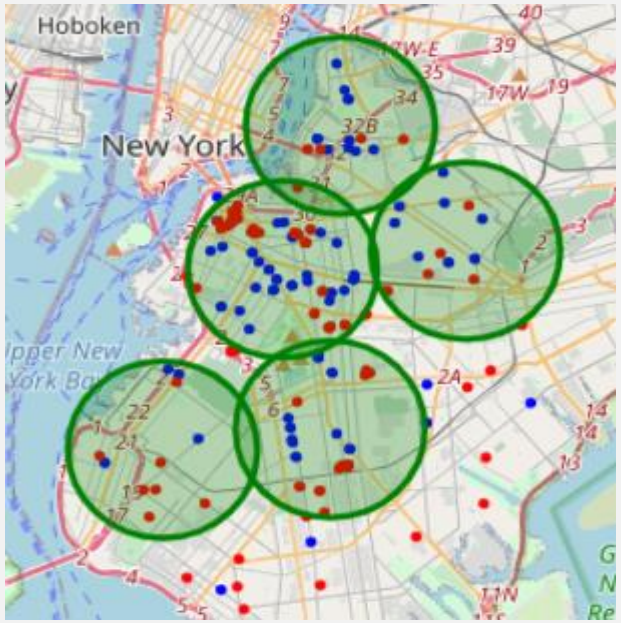
Using the **Foursquare** API service, we have located all the Indian restaurants in the New York city. In the map, we have displayed Indian restaurants for Bronx, Brooklyn, Manhattan, Queens, and Staten Island.

From the map, we can say that the stakeholders will face lot challenges in Manhattan. Staten Island and Bronx have very less Indian restaurants. We have to take other factors here. Hence, we will not consider Manhattan, Staten Island, and Bronx in our further analysis.

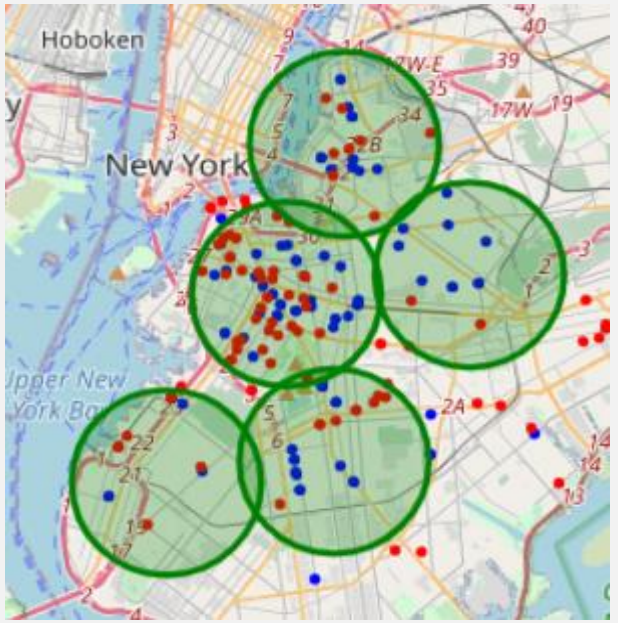
Analysis – Brooklyn



Indian Restaurants
vs
Arts & Entertainment Places



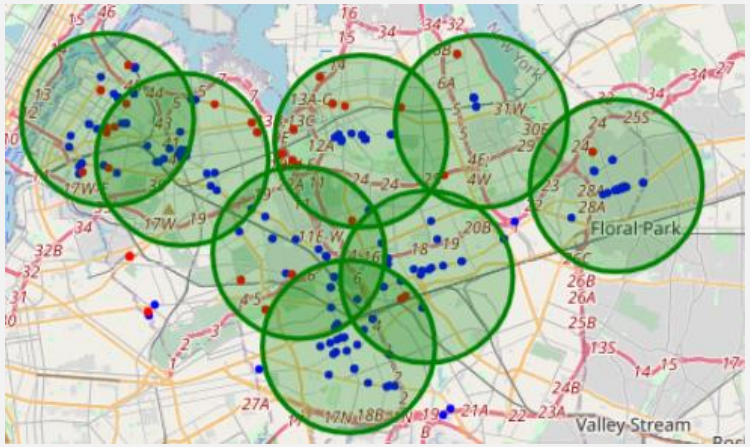
Indian Restaurants
vs
College & University



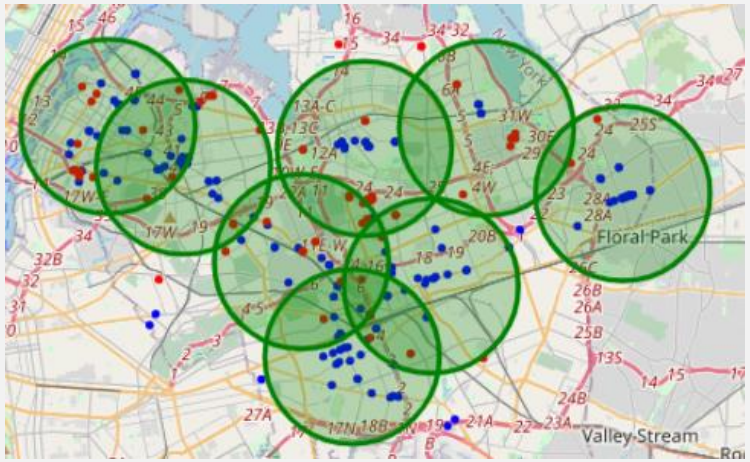
Indian Restaurants
vs
Professional & Other Places

Looks like Brooklyn has good number of restaurants. Using K-Means clustering, we get to know about the density of the Indian restaurants. We also checked the other attractive places where people like to visit frequently. This will help the stakeholder to identify suitable location.

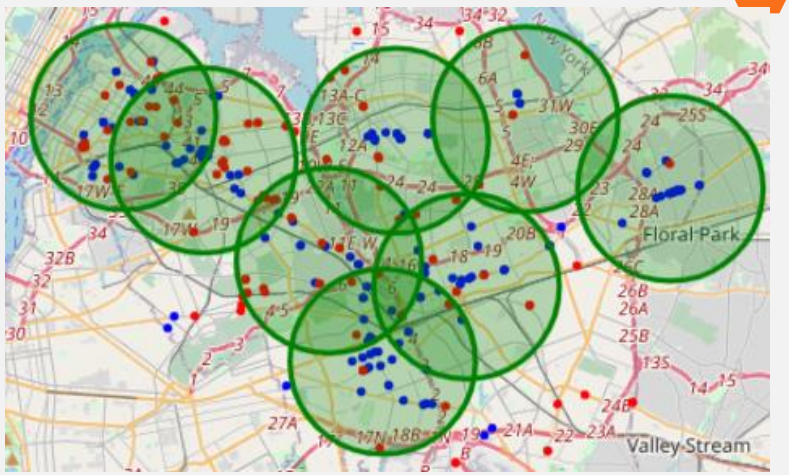
Analysis – Queens



Indian Restaurants
vs
Arts & Entertainment Places



Indian Restaurants
vs
College & University



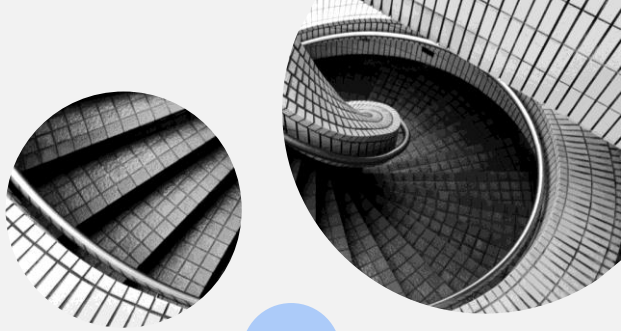
Indian Restaurants
vs
Professional & Other Places

Looks like Queens has good number of restaurants. Using K-Means clustering, we get to know about the density of the Indian restaurants. We also checked the other attractive places where people like to visit frequently. This will help the stakeholder to identify suitable location.

Results and Discussion

Let’s dive in

Results and Discussion

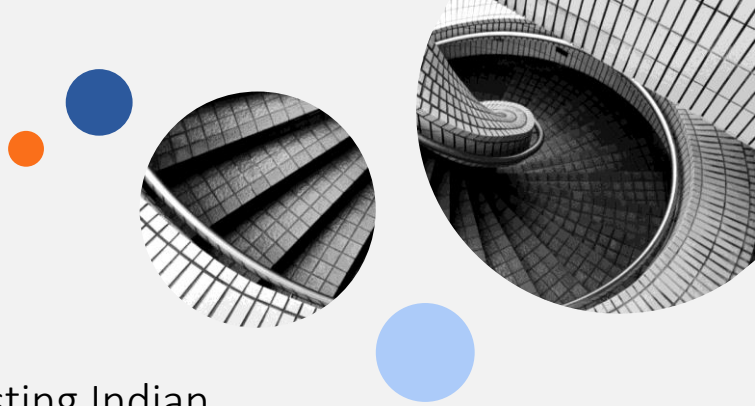


Our analysis shows that although there is a great number of Indian restaurants in New York city. Manhattan has higher number of Indian restaurants and restaurants in densed locations compared to other boroughs. Open a restaurant in Manhattan is challenging. Our stakeholder will face lot challenges. They have compete again existing Indian restaurants.
Bronx and Staten Island have very less Indian restaurants. We have to consider some other factors here before open a restaurants. So we will ignoring these places for now.

Though Brooklyn and Queen has good number of restaurants. Compared to the square area, they still have less than Manhattan. In most of the locations, Indian restaurants are not dense. Using K-Means clustering, where we get to know about the density of the Indian restaurants. To open a new restaurants, our stakeholders have the following options.

- Open a new restaurant near the dense "Art & Entertainment" category venues will give an advantage. Most of the people will visit these venues over the weekends. People will notice the new place and they would like to try it out too.
- Open a new restaurant near the dense "College & University" and "Professional & Other Places" category venues will give an advantage too. The daily commuters to these area might show interest in the People will notice the new place and they would like to try it out too.

Results and Discussion (Contd..)



Purpose of this analysis was to only provide info on areas that are not crowded with existing Indian restaurants. It is also possible that there is a very good reason for small number of restaurants in any of those areas, reasons which would make them unsuitable for a new restaurant regardless of lack of competition in the area. Recommended zones should therefore be considered only as a starting point for more detailed analysis which could eventually result in location which has not only no nearby competition but also other factors taken into account and all other relevant conditions met.

Conclusion

Purpose of this project was to identify a location for a new Indian restaurant in order to help our stakeholders in narrowing down the search for optimal location for a new Indian restaurant. By finding restaurant density from Foursquare data, we have identified boroughs that justify further analysis (Brooklyn and Queen). Clustering helped use to understand about the areas of boroughs. We also looked for other important category dense areas where Indian restaurants are less and our stakeholders have good opportunity to open a new restaurants.

Final decision on restaurant location will be made by stakeholders based on characteristics of neighborhood and taking additional factors like attractiveness of each location (proximity to park or water), levels of noise / proximity to major roads, real estate availability, prices, social and economic dynamics of every neighborhood etc.

Sunil Kunchoor Basavaraju