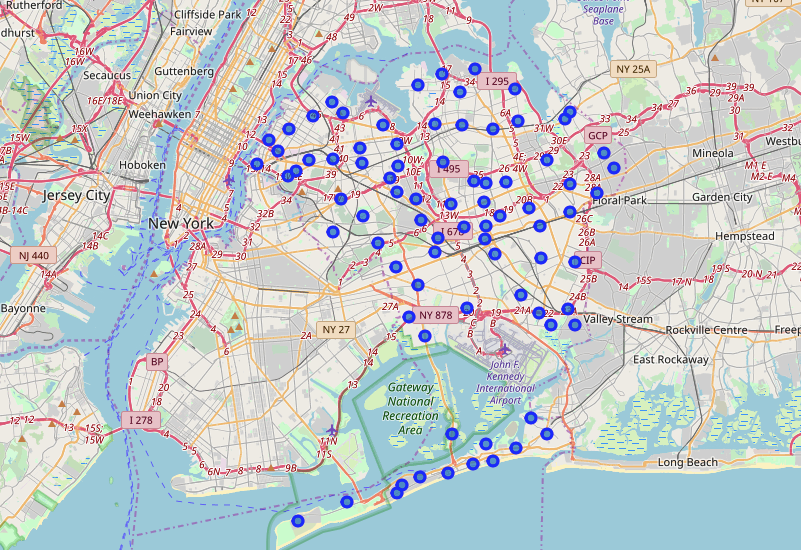
**Finding the Best Location to Open a Bar in Queens, NY**

1. Introduction

The borough of Queens in New York City is one of the most diverse and dense boroughs in the east coast. Finding a good location to open a bar in Queens can be a vital step when it comes to be a successful business. Python can be used to handle this type of analysis because it has a lot of useful libraries that can combine different things into meaningful data. Therefore, the task of finding a good location can become less complex and time consuming by using a good data set and the appropriate algorithms.

1. Data Extraction

The Queens neighborhood information such as names, longitude, and latitude comes from the JSON file: *nyu\_2451\_34572-geojson.json*. The JSON file is transformed into a pandas dataframe containing the following columns: Borough, Neighborhood, Latitude and Longitude. The venue information of each neighborhood is obtained through the Foursquare API. The goal is to find the top 10 venue categories by neighborhood.

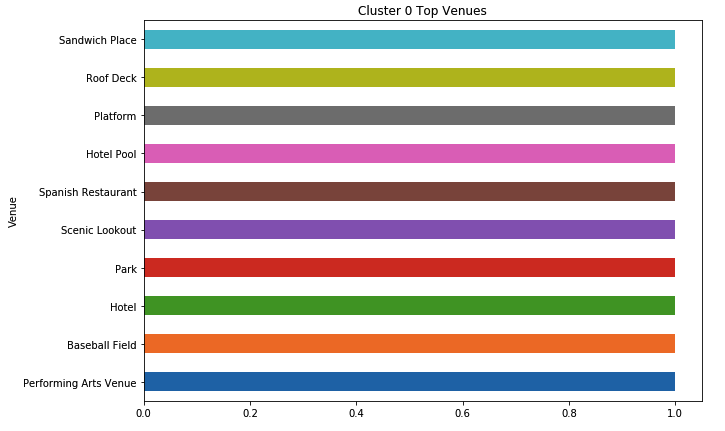


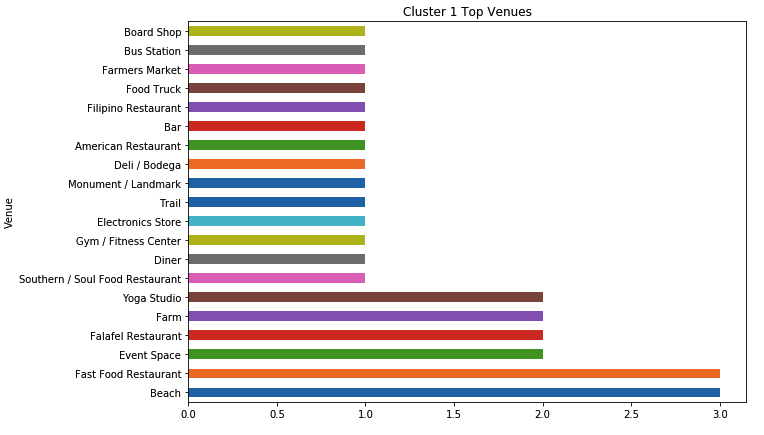
1. Methodology

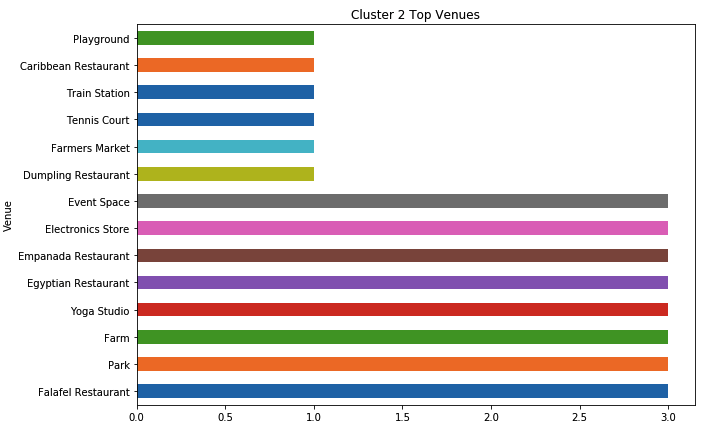
By exploring the most popular venues in each neighborhood, we can determine which neighborhoods are more likely to have less bars. After extracting the venues information from Foursquare, a dataframe is created by grouping venues by category.

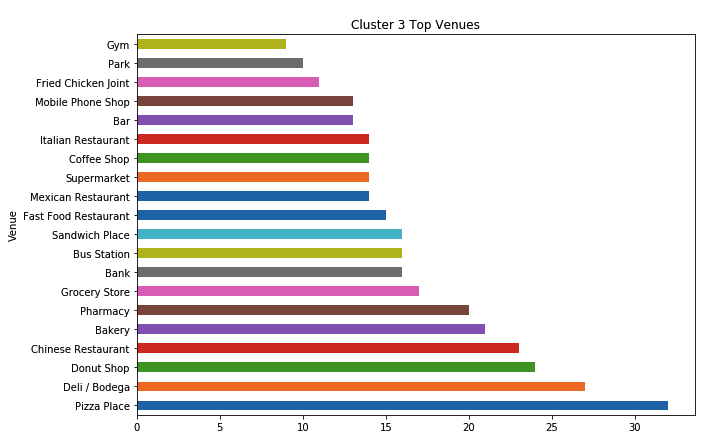
* 1. Clustering Analysis

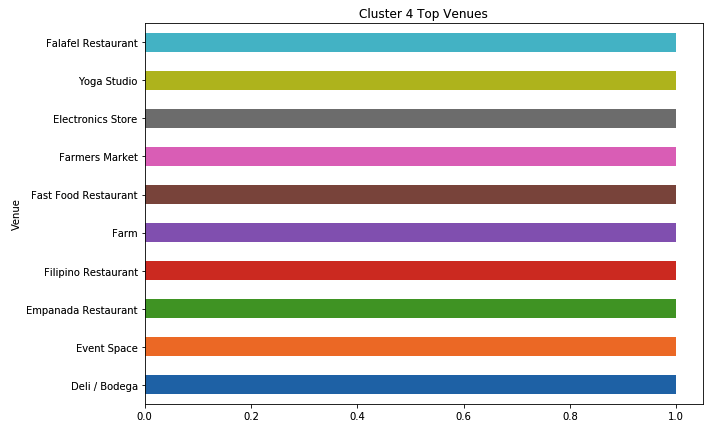
A clustering analysis is used to determine the relationships among the neighborhoods based on the venue categories. Before the analysis, one hot encoding is performed on the Foursquare dataframe and the mean per neighborhood is calculated. By using the k-means clustering algorithm, 5 clusters are created. A column containing the cluster number (0-4) is added to our final dataframe, which contains the borough name, neighborhood name, latitude, longitude, and top 10 venues.





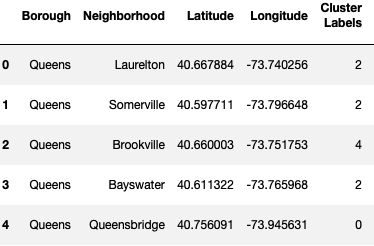






1. Analysis Results

By looking at each cluster, we can see that clusters 1 and 3 contain “bar” as a top venue, so we can discard the neighborhoods in cluster 1 and 3. Consequently, the following neighborhoods can be a good location to open a bar:



1. Observations

Bayswater or Somerville can be a perfect location if the plan is to open a bar in the summer because they are close to the beach. On the other hand, Laurelton and Brookville are close to the JFK Airport, which can attract travelers. In addition, information containing the average income of the final 5 neighborhoods can help decide which neighborhood is the best to open a bar.

1. Conclusion

A good data set and the right analysis can speed up the process of finding a good place to start a business. Foursquare makes things easier by providing existing business information. By using Python, we can combine information from multiple sources, such as files and APIs, and use different analysis tools or algorithms depending on the needs of the project.