



FINDING THE BEST LOCATION TO OPEN A BAR IN QUEENS, NY

Introduction

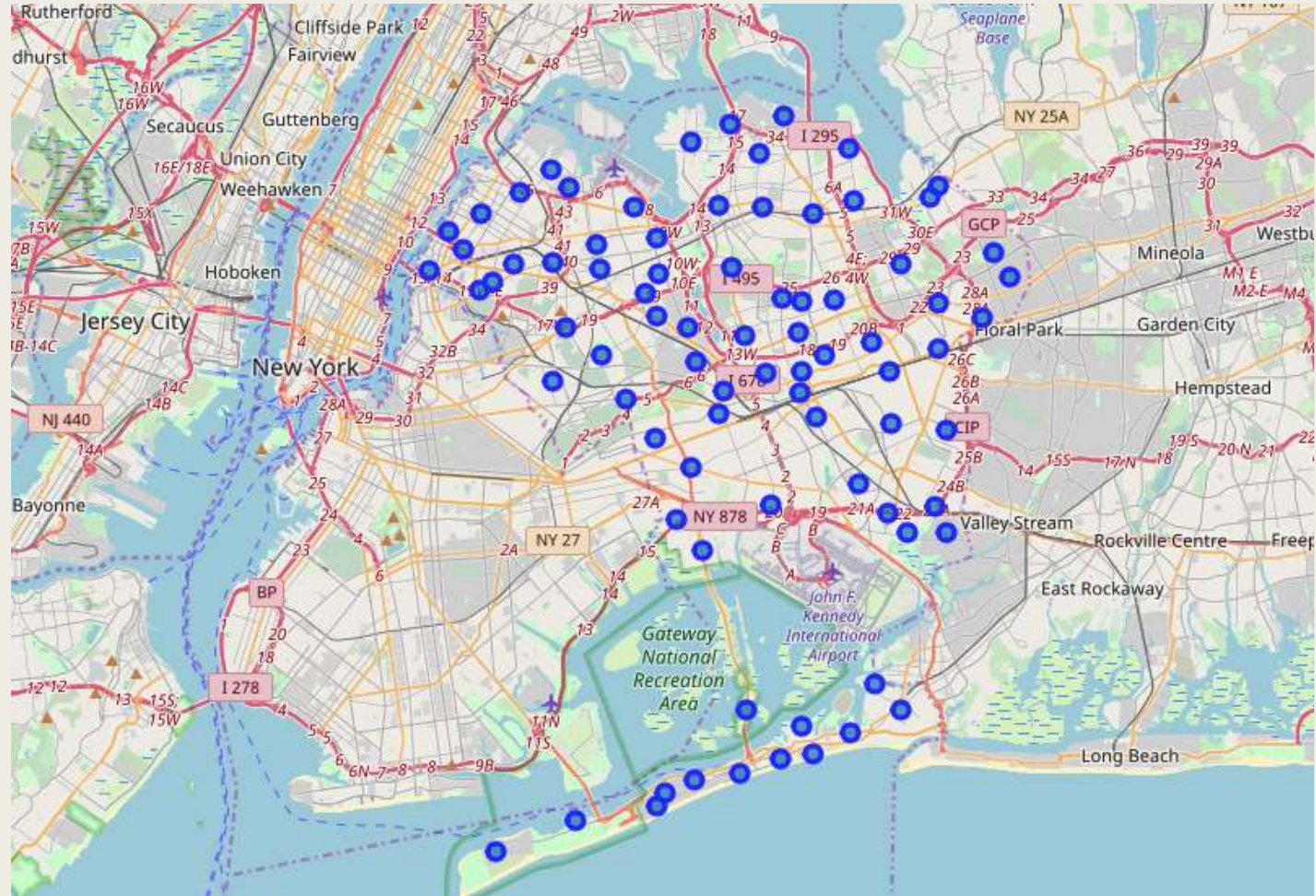
- The borough of Queens in New York City is one of the most diverse and dense boroughs in the east coast
- Finding a good location to open a bar in Queens can be a vital step when it comes to be a successful business
- By using a good data set and the appropriate algorithms, the task of finding a good location can become less complex and time consuming

Data Extraction

- Queens neighborhood names, latitude, and longitude are extracted from [nyu_2451_34572-geojson.json](#)
- Venues information are extracted from *Foursquare* through their API

	Borough	Neighborhood	Latitude	Longitude
0	Queens	Astoria	40.768509	-73.915654
1	Queens	Woodside	40.746349	-73.901842
2	Queens	Jackson Heights	40.751981	-73.882821
3	Queens	Elmhurst	40.744049	-73.881656
4	Queens	Howard Beach	40.654225	-73.838138

Neighborhoods in Queens, NY



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 can be created by grouping venues by category
- Finally, a clustering analysis is performed on the dataframe containing the venue categories and neighborhoods
- Based on the analysis, we can discard the neighborhoods that have “*bar*” as a popular venue

Foursquare Data

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Arverne	Surf Spot	Metro Station	Playground	Beach	Sandwich Place	Bed & Breakfast	Thai Restaurant	Coffee Shop	Donut Shop	Board Shop
1	Astoria	Middle Eastern Restaurant	Bar	Hookah Bar	Greek Restaurant	Seafood Restaurant	Bakery	Italian Restaurant	Latin American Restaurant	Salon / Barbershop	Coffee Shop
2	Astoria Heights	Pizza Place	Bakery	Bus Line	Burger Joint	Deli / Bodega	Shopping Mall	Bowling Alley	Supermarket	Moving Target	Chinese Restaurant
3	Auburndale	Italian Restaurant	Hookah Bar	Discount Store	Fast Food Restaurant	Gymnastics Gym	Toy / Game Store	Korean Restaurant	Train	Athletics & Sports	Supermarket
4	Bay Terrace	Clothing Store	Women's Store	American Restaurant	Kids Store	Mobile Phone Shop	Donut Shop	Lingerie Store	Cosmetics Shop	Steakhouse	Men's Store

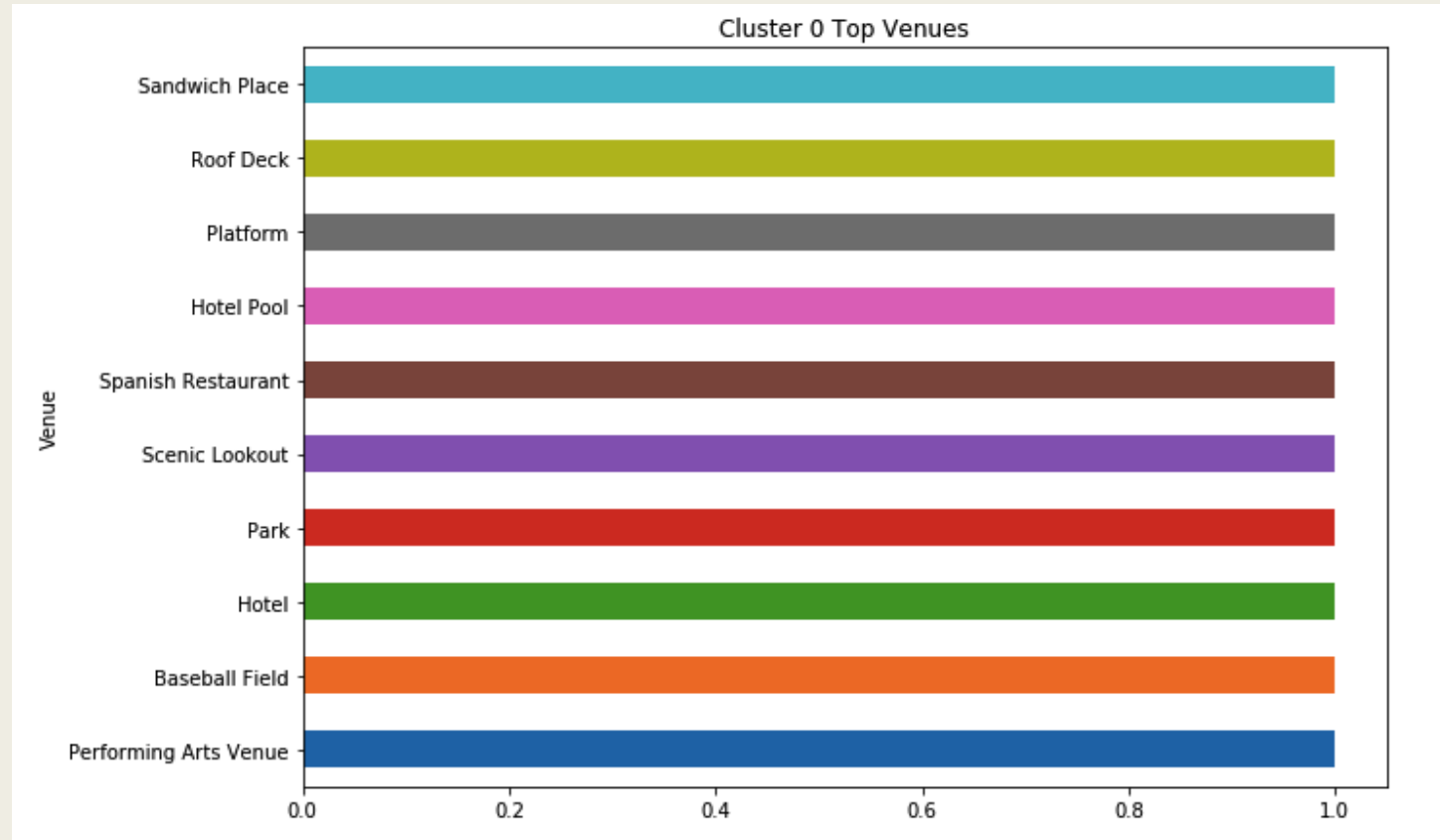
Clustering Analysis

- The algorithm used is k-means clustering
- Number of clusters: 5
- Before the analysis, one hot encoding is performed on the Foursquare dataframe and the mean per neighborhood is calculated

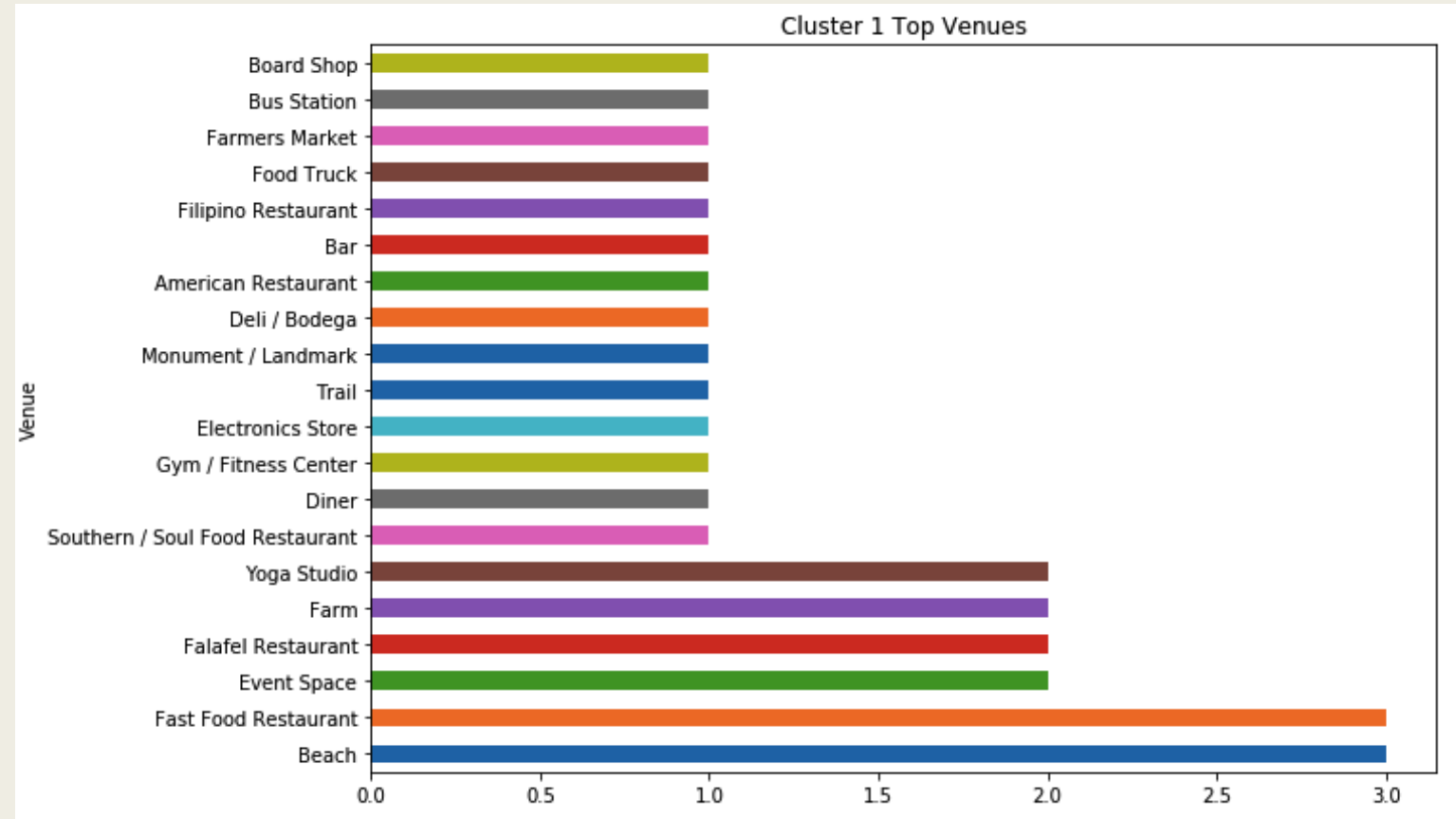
Clustering Data Set

	Neighborhood	Accessories Store	Afghan Restaurant	Airport Terminal	American Restaurant	Antique Shop	Arepa Restaurant	Argentinian Restaurant	Art Gallery	Art Museum	...	Video Game Store	Video Store	Vietnamese Restaurant
0	Arverne	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000	0.0	0.0	...	0.000000	0.000000	0.000000
1	Astoria	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000	0.0	0.0	...	0.000000	0.000000	0.000000
2	Astoria Heights	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000	0.0	0.0	...	0.000000	0.000000	0.000000
3	Auburndale	0.000000	0.000000	0.000000	0.052632	0.000000	0.000000	0.000	0.0	0.0	...	0.000000	0.000000	0.000000
4	Bay Terrace	0.024390	0.000000	0.000000	0.048780	0.000000	0.000000	0.000	0.0	0.0	...	0.024390	0.000000	0.000000
5	Bayside	0.000000	0.000000	0.000000	0.042254	0.000000	0.000000	0.000	0.0	0.0	...	0.000000	0.000000	0.000000
6	Bayswater	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000	0.0	0.0	...	0.000000	0.000000	0.000000
7	Beechhurst	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000	0.0	0.0	...	0.000000	0.000000	0.000000
8	Bellaire	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000	0.0	0.0	...	0.000000	0.000000	0.000000

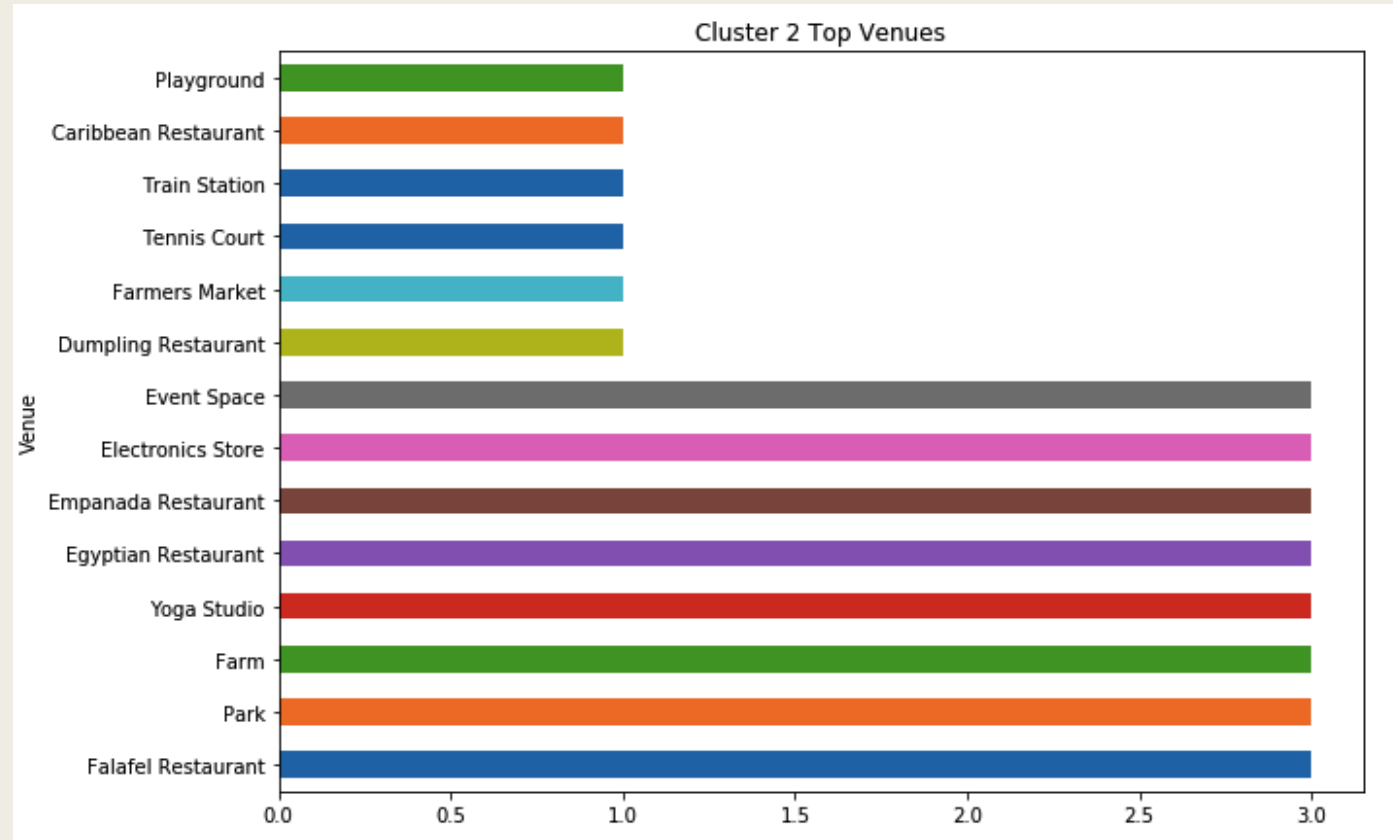
First Cluster



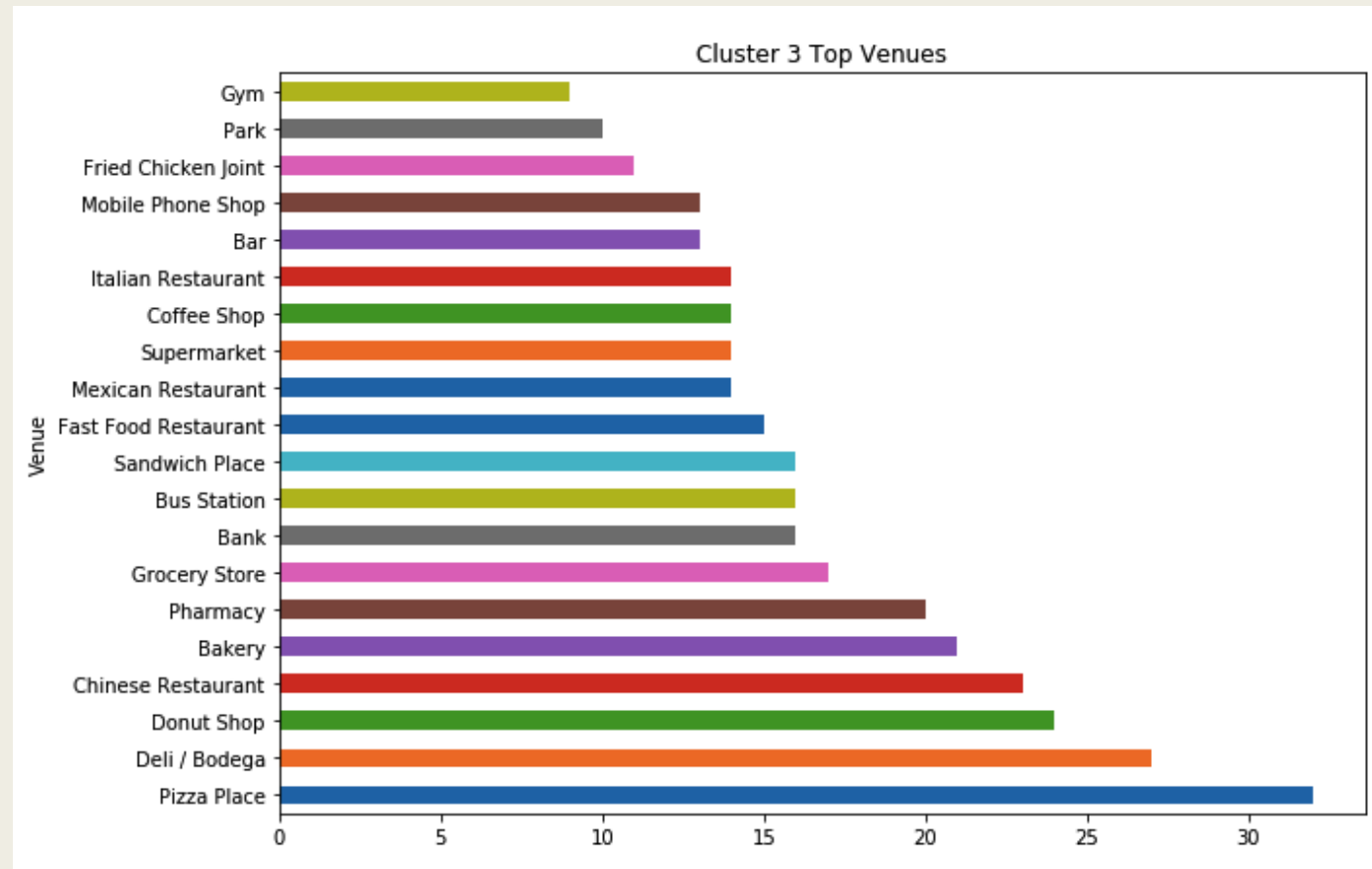
Second Cluster



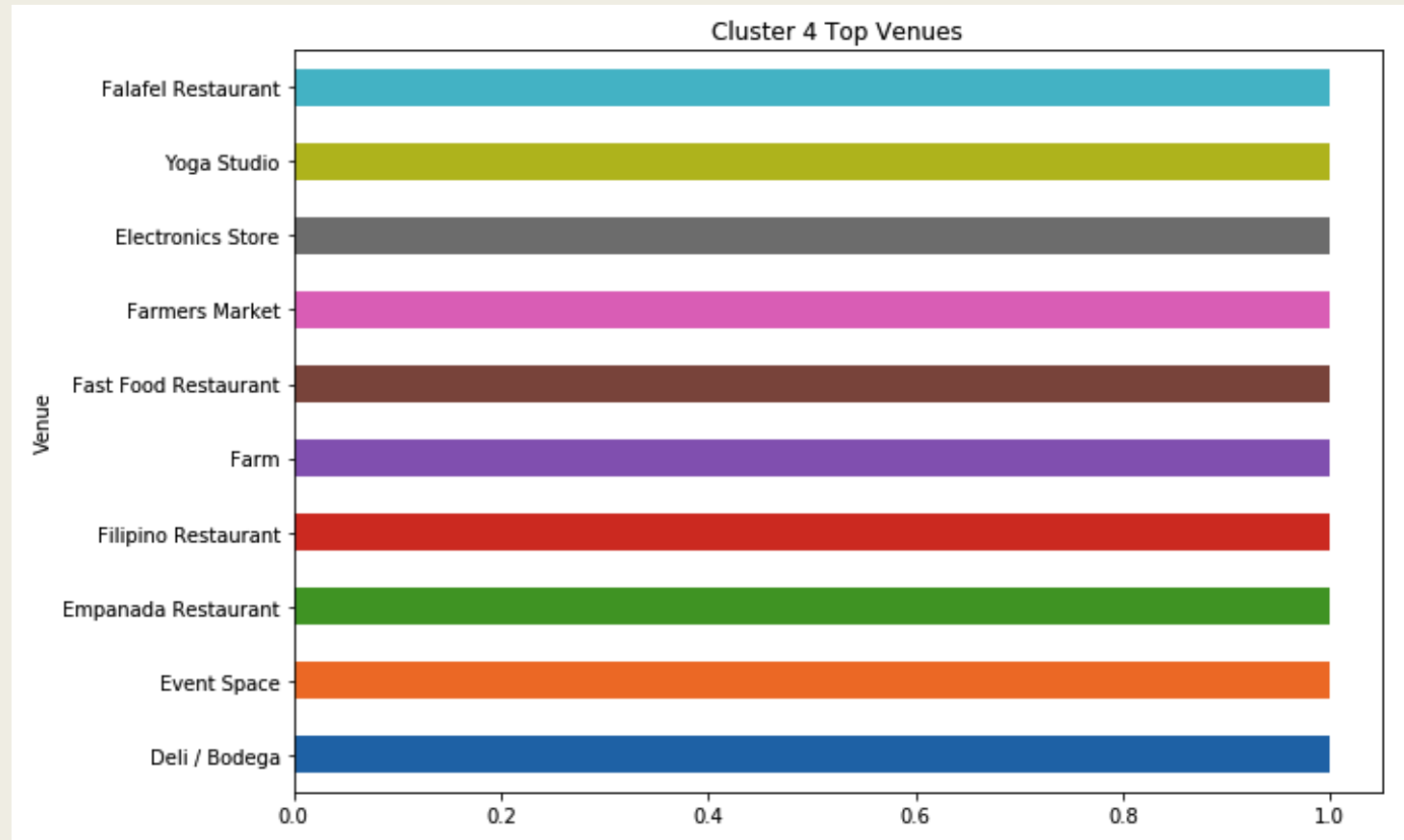
Third Cluster



Fourth Cluster



Fifth Cluster



Analysis Results

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

	Borough	Neighborhood	Latitude	Longitude	Cluster Labels
0	Queens	Laurelton	40.667884	-73.740256	2
1	Queens	Somerville	40.597711	-73.796648	2
2	Queens	Brookville	40.660003	-73.751753	4
3	Queens	Bayswater	40.611322	-73.765968	2
4	Queens	Queensbridge	40.756091	-73.945631	0

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

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