

Best Cluster of Neighborhoods from Queens and Manhattan Boroughs Offering Multi-cultural Food Restaurants.

Pravinkumar Kandhare, PhD Student.
University of Alabama, Birmingham

Contents:

- 1. Problem Statement, Audience and Dataset**
- 2. Visualization of neighborhoods on map**
- 3. Manhattan and Queens restaurant venue dataset from Foursquare API**
- 4. K-means clustering algorithm**
- 5. Analyze clusters of Manhattan and Queens**
- 6. Best cluster offering a greater number of multi-cultural restaurants**
- 7. Conclusion**

1. Problem Statement , Audience and Dataset:

Problem Statement:

*“Compare the boroughs (Manhattan and Queens) of the New York city and determine the **best cluster** that offers large number of multi-cultural food restaurants.”*

Clustering neighborhood of each borough will be done by choosing top three most common venue restaurants categories. Then, analyzing all clusters we would like to ask following questions.

- Which borough cluster has largest variety of multi-cultural restaurants.
- Discussion on cluster distribution on each borough

Who is Audience:

Multicultural Food Lovers

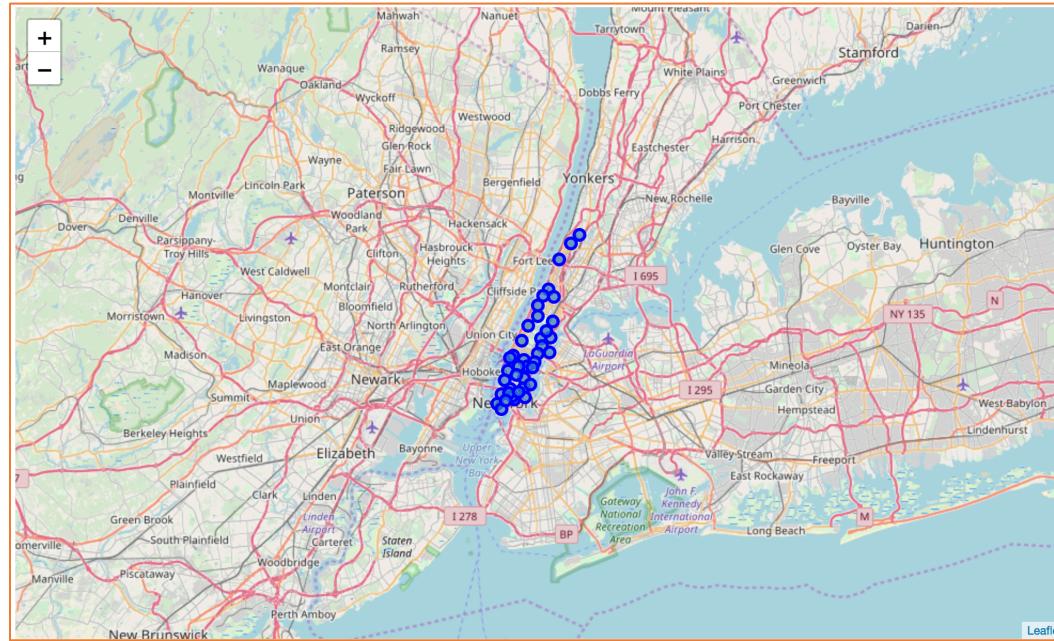
Datasets:

1. New York city dataset: contain boroughs and their neighborhood along with latitude and longitude features. (“https://cocl.us/new_york_dataset”)
2. Foursquare venue dataset: contains venue, neighborhood and other venue related information. (“<https://api.foursquare.com>”)

2. Visualization of neighborhoods on Maps:

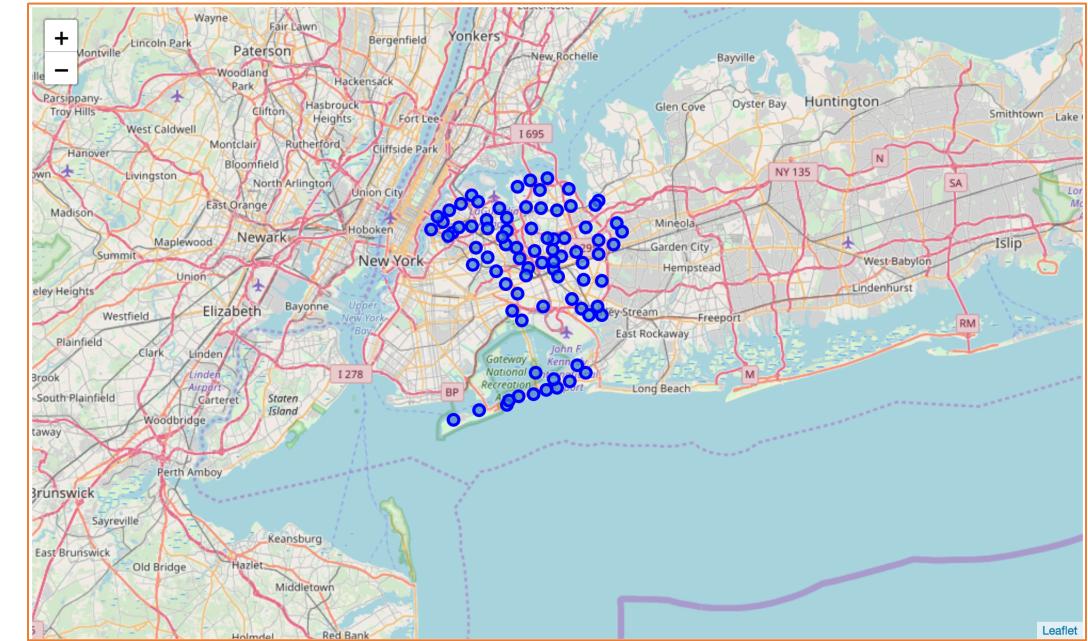
Manhattan Neighborhoods on Map

Total Number of Manhattan's Neighborhoods : 39



Queens Neighborhoods on Map

Total Number Queen's Neighborhoods: 70



New York dataset processed for getting two different dataset (Manhattan and Queens)

Neighborhood of each borough is displayed on Map

Queen has **more number** of neighborhoods **compared** to Manhattan

3. Manhattan and Queens restaurant venue datasets from Foursquare API:

Our interest is in only top 3 most common venue restaurants of each neighborhood among the Manhattan and Queens borough's dataset.

Filtered sample restaurant datasets for Manhattan and Queens are as follows.

Manhattan Total Neighborhoods : 39

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
0	Battery Park City	Italian Restaurant	Sushi Restaurant	American Restaurant
1	Carnege Hill	French Restaurant	Japanese Restaurant	Indian Restaurant
2	Central Harlem	African Restaurant	American Restaurant	Seafood Restaurant
3	Chelsea	Italian Restaurant	American Restaurant	Seafood Restaurant
4	Chinatown	Chinese Restaurant	Vietnamese Restaurant	American Restaurant

Queens Total Neighborhoods: 70

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
0	Arverne	Thai Restaurant	Vietnamese Restaurant	Dumpling Restaurant
1	Astoria	Middle Eastern Restaurant	Greek Restaurant	Seafood Restaurant
2	Astoria Heights	Italian Restaurant	Indonesian Restaurant	Hotpot Restaurant
3	Auburndale	Italian Restaurant	American Restaurant	Korean Restaurant
4	Bay Terrace	American Restaurant	Gluten-free Restaurant	Vietnamese Restaurant

Top 3 most common restaurant venues in each neighborhood of Manhattan and Queens

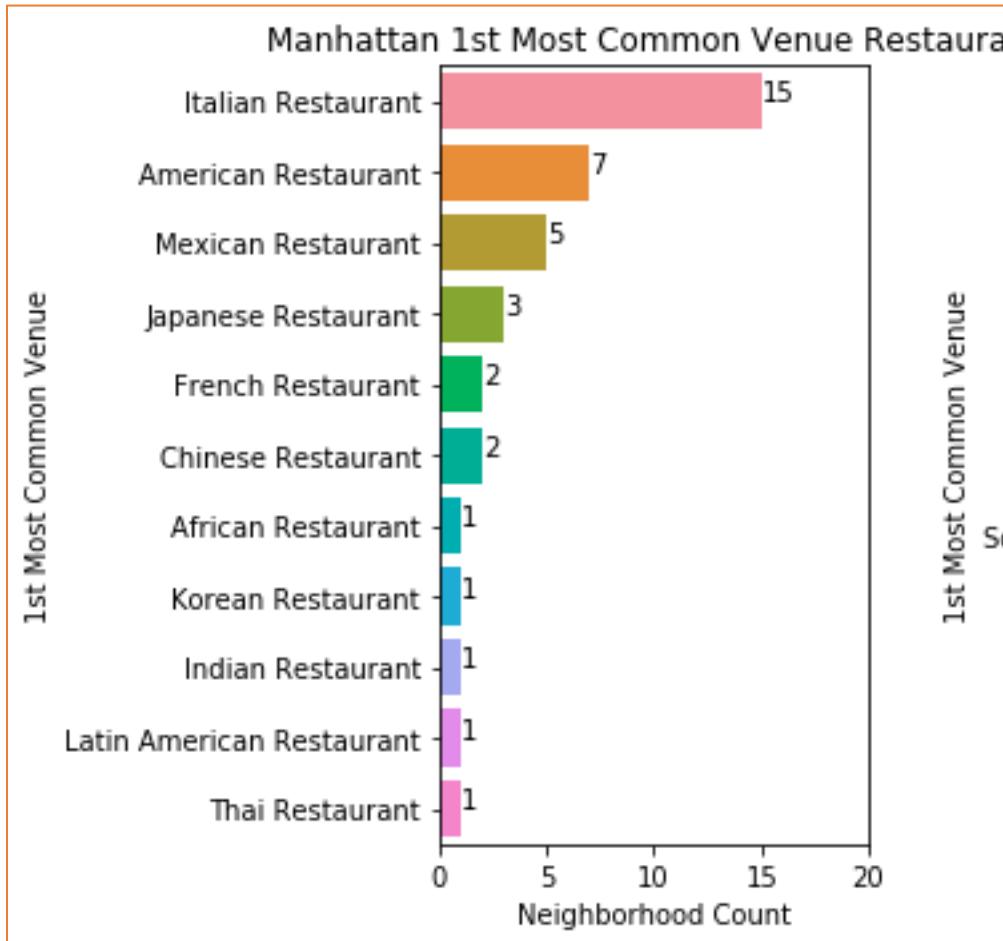
Note*: Only first five values of datasets are displayed

Let's analyze top 3 most common venue restaurants of Manhattan and Queens and check their popular (highest count) restaurant category.

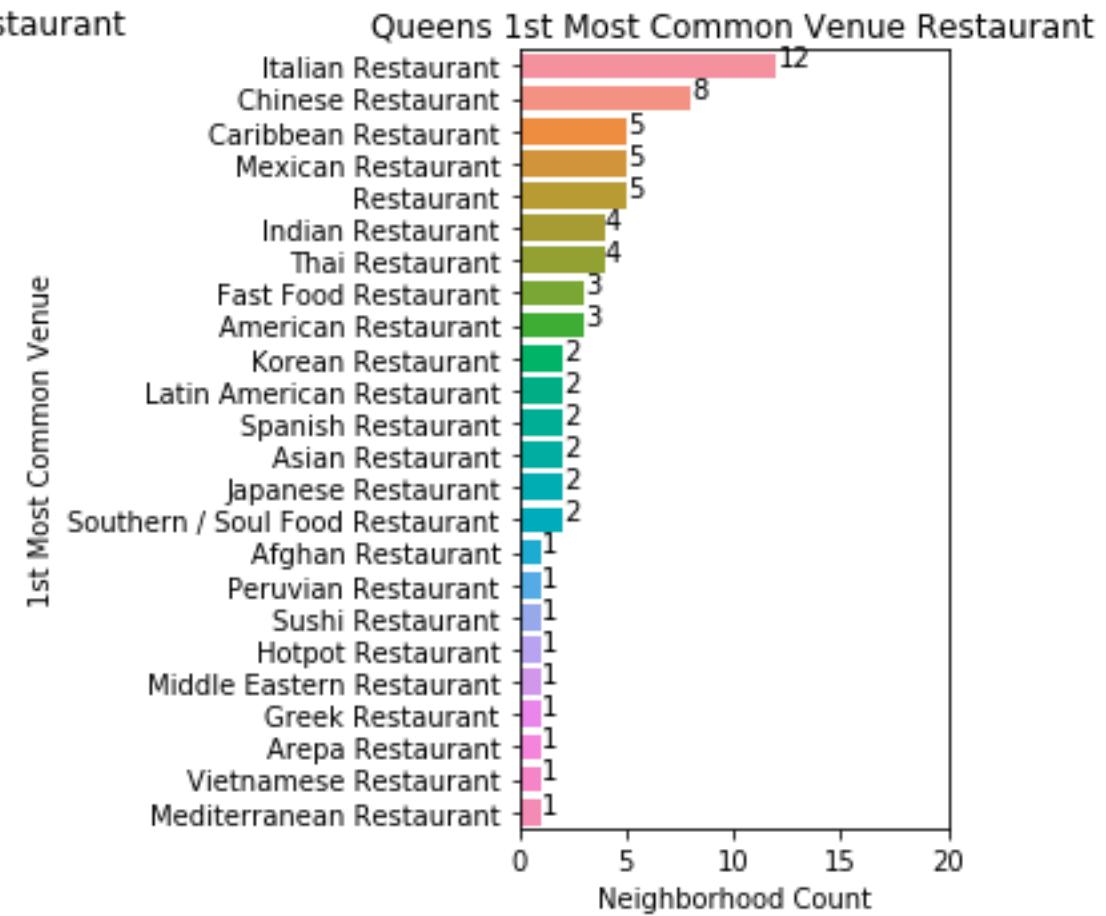
1. 1st Most Common Restaurant Venue Category of Manhattan and Queens
2. 2nd Most Common Restaurant Venue Category of Manhattan and Queens
3. 3rd Most Common Restaurant Venue Category of Manhattan and Queens

3.1 1st Most Common Restaurant Venue Category of Manhattan and Queens:

Manhattan Total Neighborhoods : 39



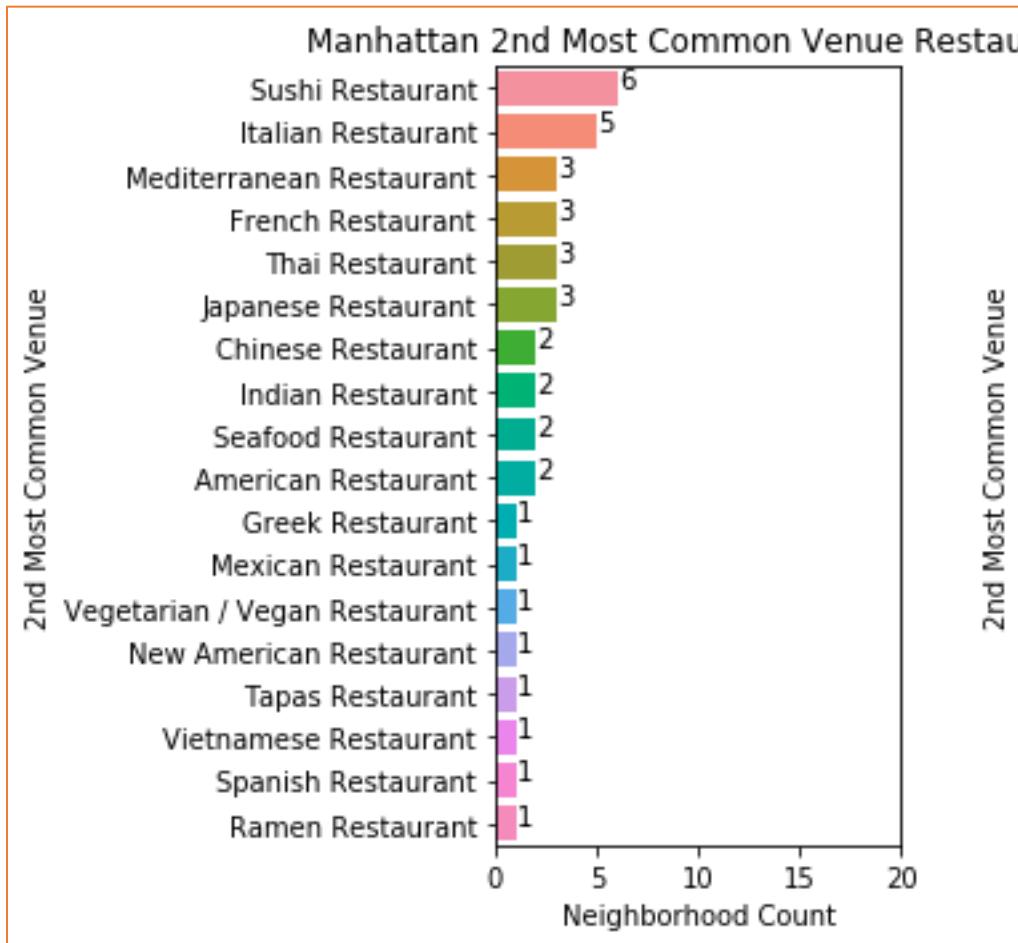
Queens Total Neighborhoods: 70



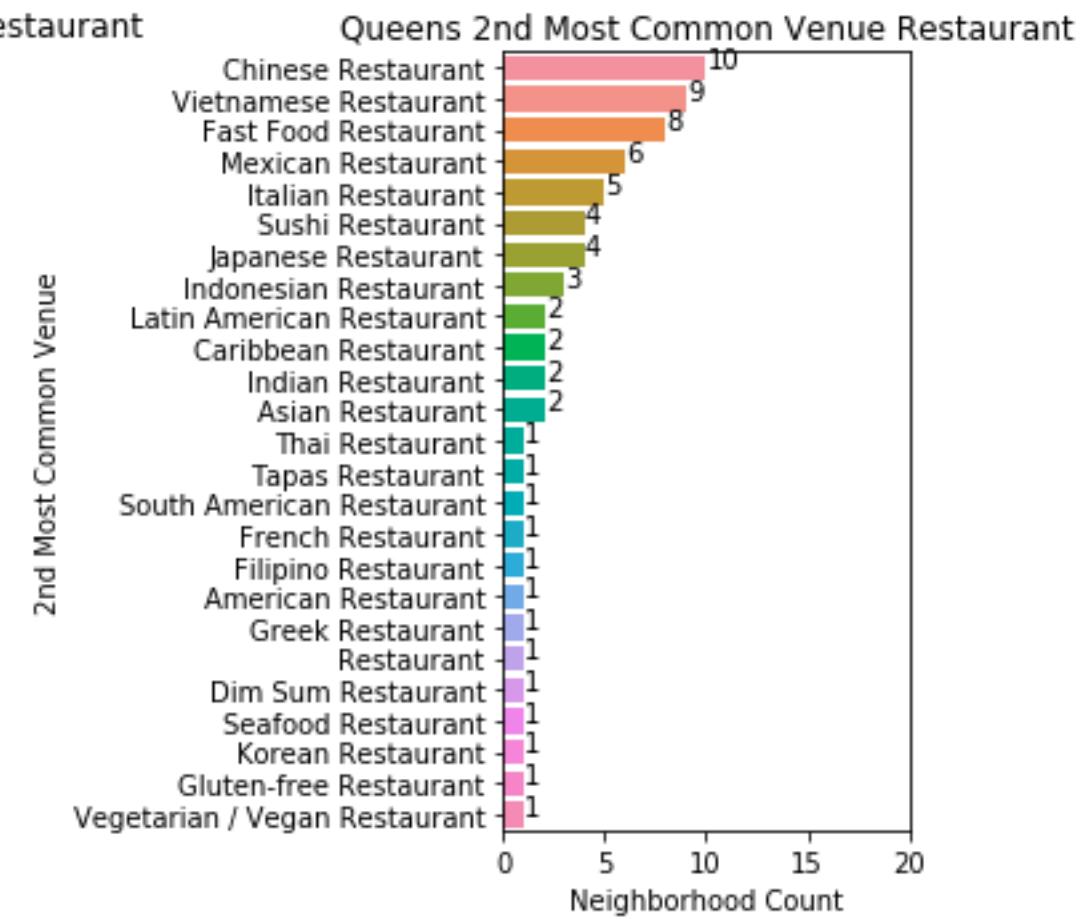
Queen has a greater number of “1st most common venue restaurant” categories compared to Manhattan
Italian Restaurant is a “1st most common venue restaurant” categories for Manhattan and Queens boroughs

3.2 2nd Most Common Restaurant Venue Category of Manhattan and Queens:

Manhattan Total Neighborhoods : 39



Queens Total Neighborhoods: 70

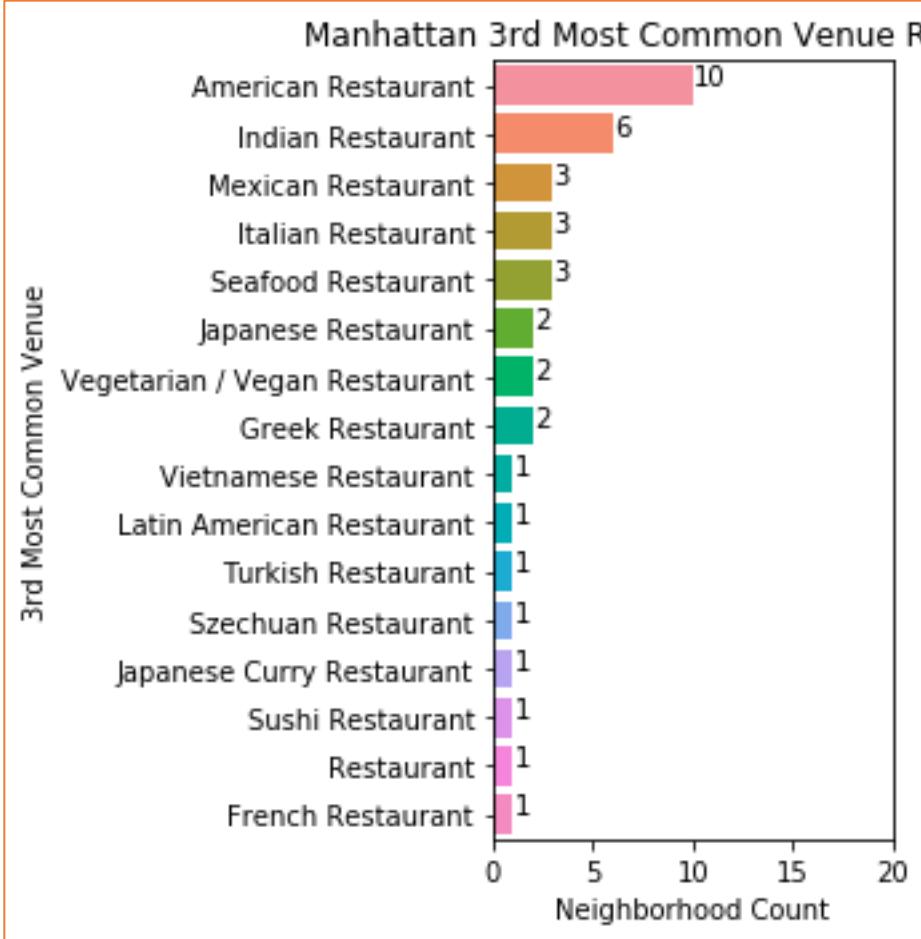


Queen has a greater number of “2nd most common venue restaurant” categories compared to Manhattan

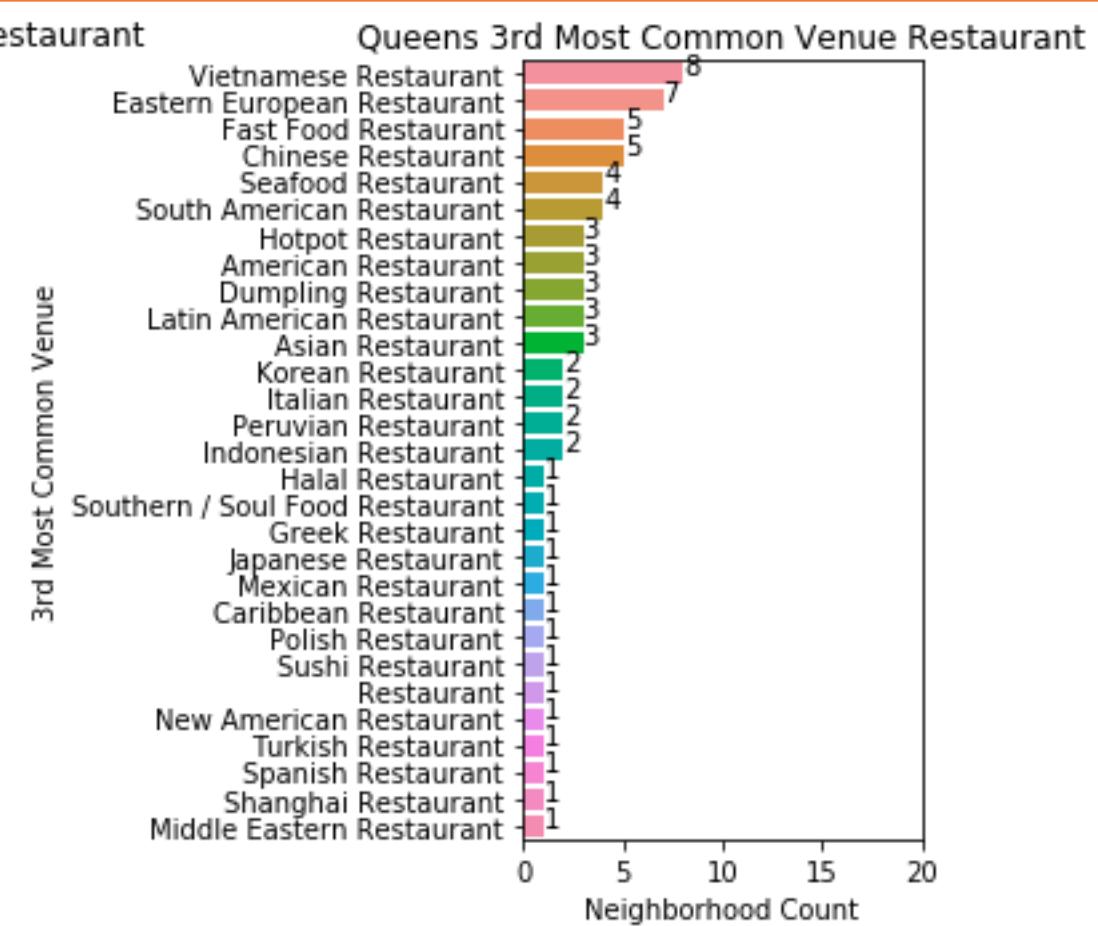
Sushi Restaurant and Chinese Restaurant are “2nd most common venue restaurant” categories for Manhattan and Queens respectively.

3.3. 3rd Most Common Restaurant Venue Category of Manhattan and Queens:

Manhattan Total Neighborhoods : 39



Queens Total Neighborhoods: 70



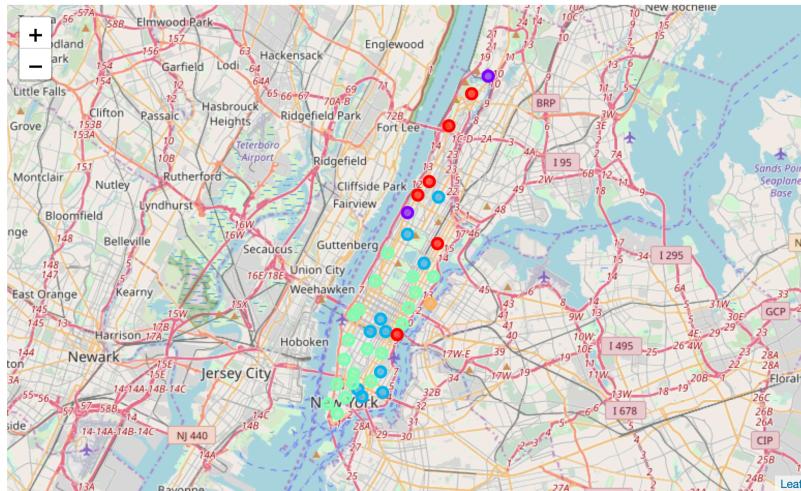
Queen has a greater number of “3rd most common venue restaurant” categories compared to Manhattan

American Restaurant & Vietnamese Restaurant are “3rd most common venue restaurant” categories for Manhattan and Queens respectively.

4. K-means clustering algorithm:

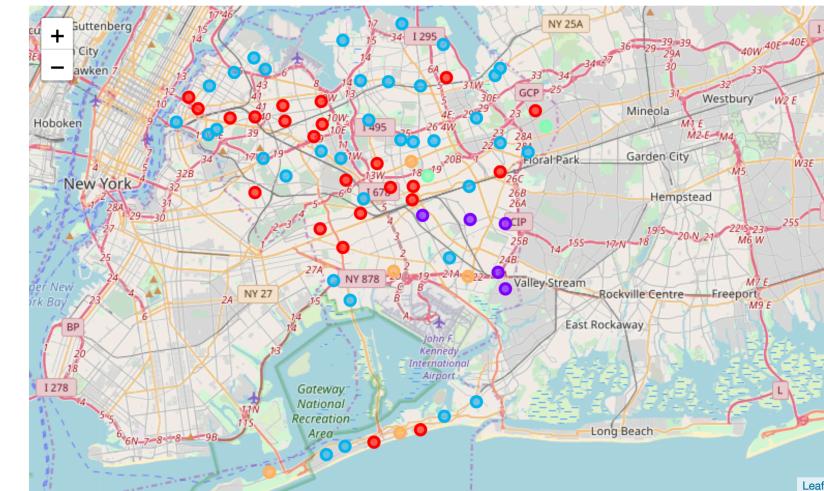
- Applied **K-mean** clustering algorithm using cluster value **K=5**.
 - Clustering results of Manhattan and Queens “top 3 most common venue restaurant” datasets.

Manhattan Neighborhoods clustering Map



Manhattan Total Neighborhood : 39

Queens Neighborhoods clustering Map



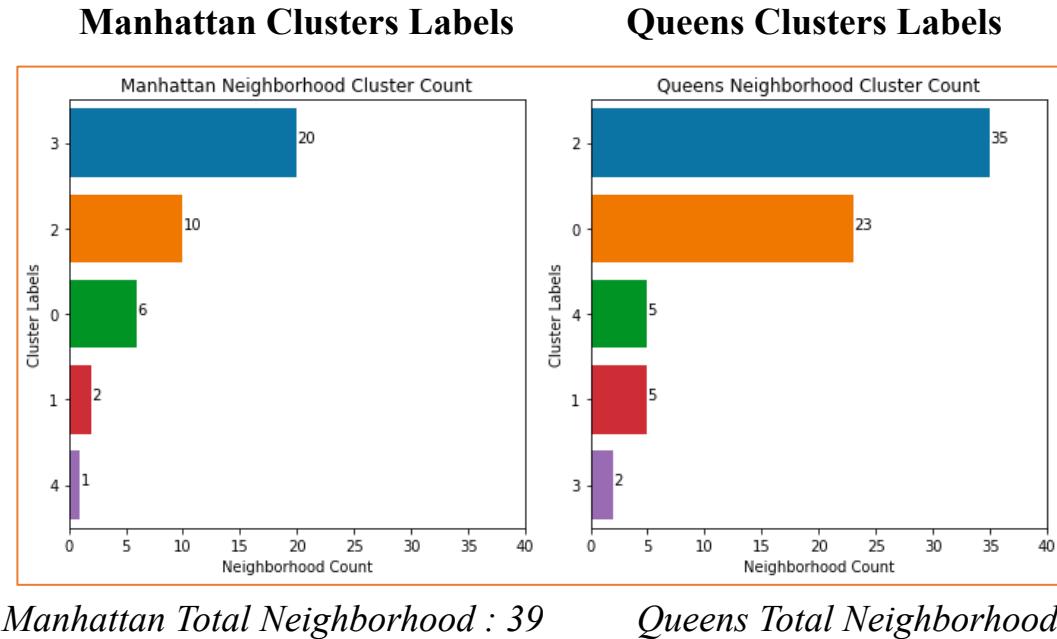
Queens Total Neighborhood: 70

Observations:

1. In Manhattan's map, "light green" colored cluster contains **greater number of neighborhoods** compared to other clusters.
 2. In Queens map, "light blue" colored cluster contains **greater number of neighborhoods** compared to other clusters
Let's check assigned label counts for each cluster.

5. Analyze clusters of Manhattan and Queens:

- Distribution of assigned clustered labels of Manhattan and Queens neighborhoods dataset .



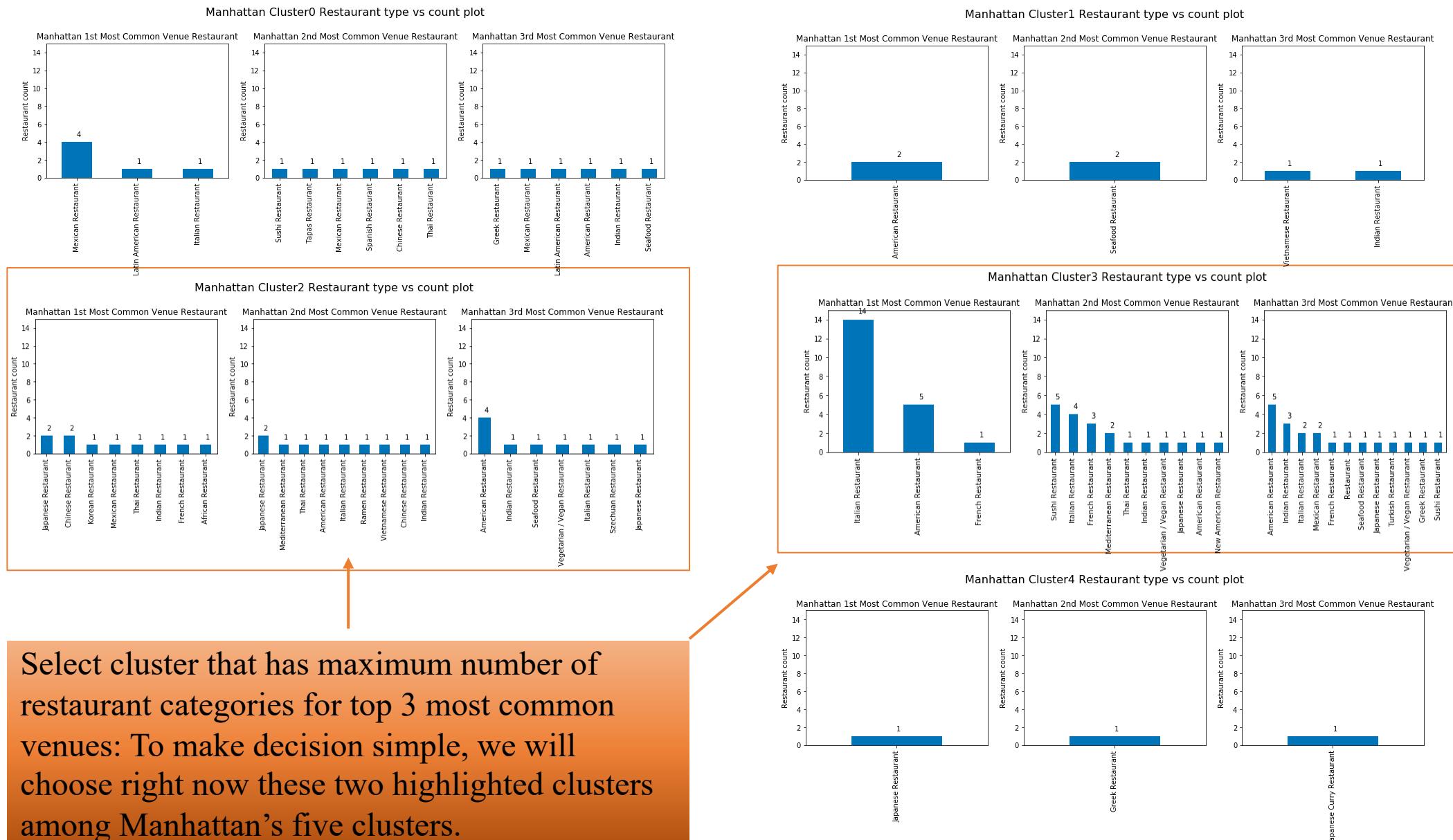
Observations:

- Manhattan **cluster number 3** contains **greater number of neighborhoods** compared to others
- Queens **cluster number 2** contains **greater number of neighborhoods** compared to others

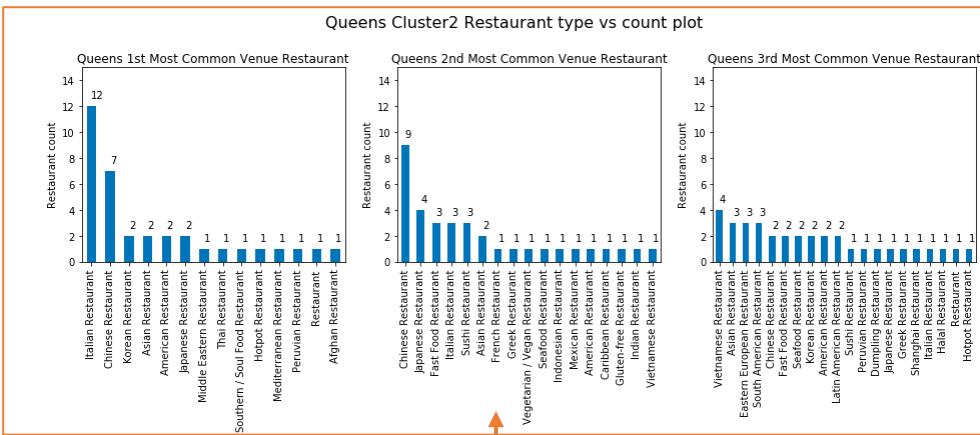
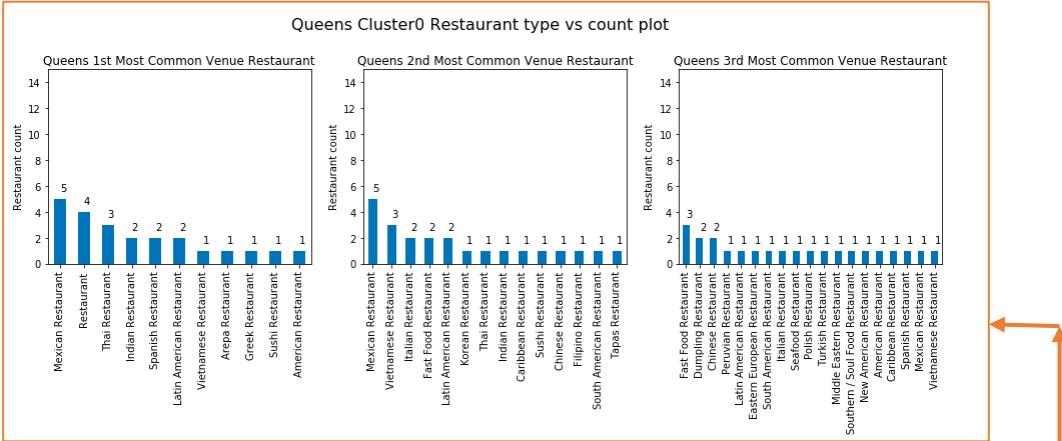
Guessing solution of our problem:

- Cluster number 2** of Queens has largest number of neighborhoods among all clusters.
- Guess can be made that **cluster 2 may offer more number of multi-cultural food restaurants**. Again, It just a guess. Let's see the correct results !!!

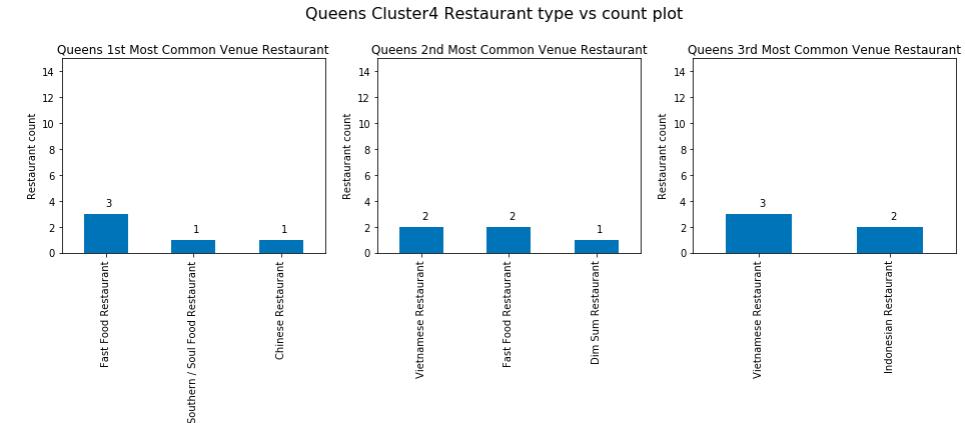
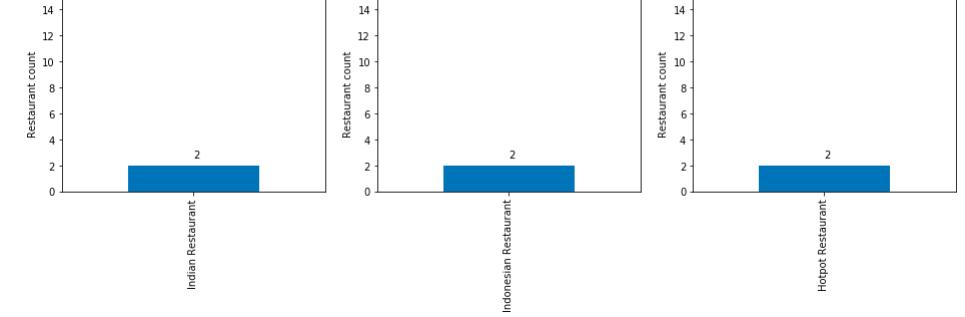
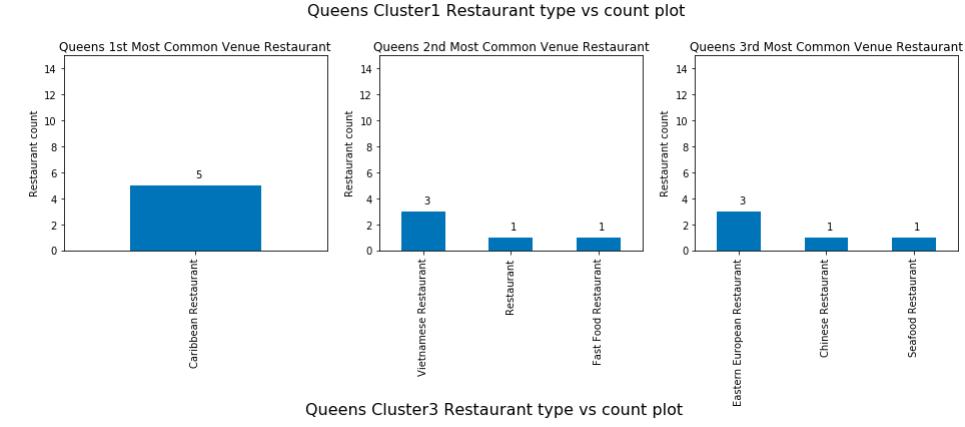
5.1 Analyze Manhattan's all five clusters along with their restaurant category:



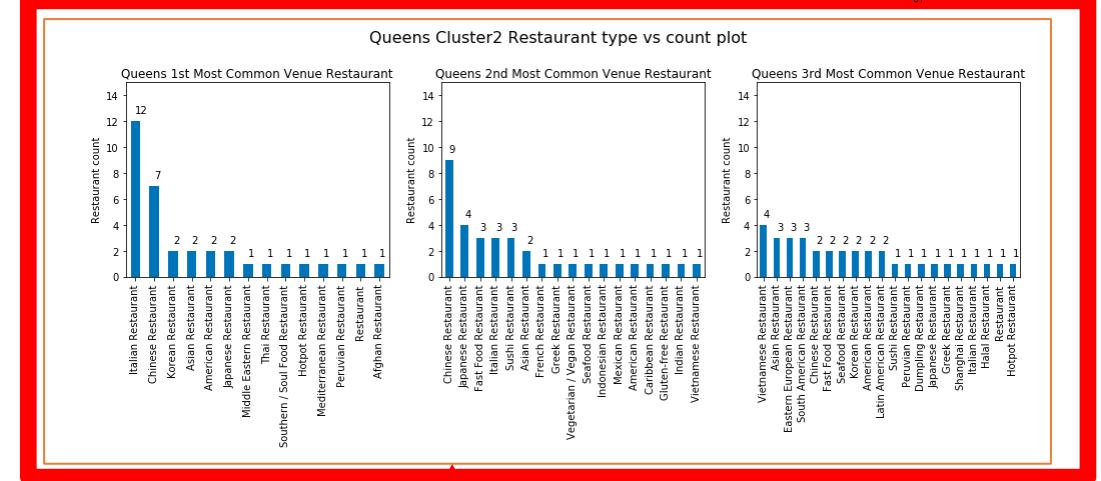
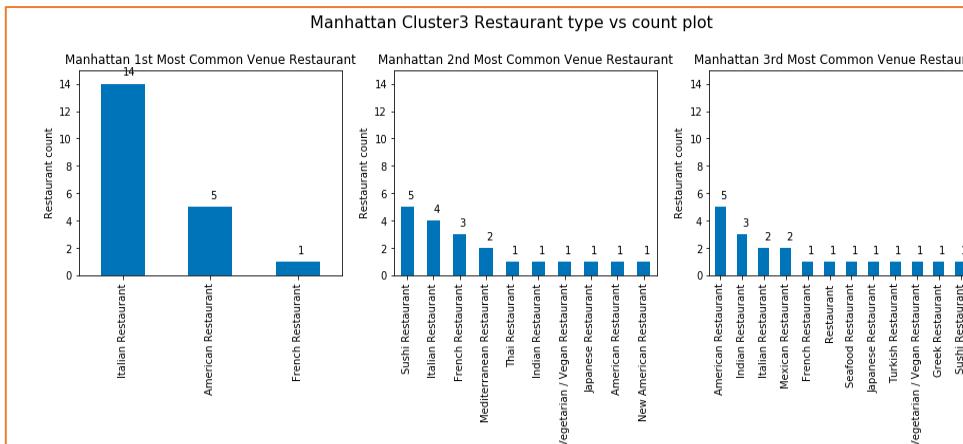
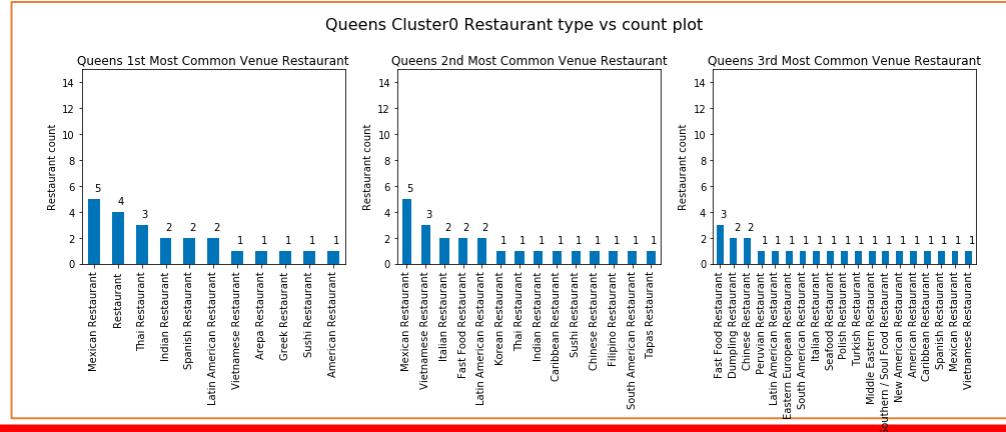
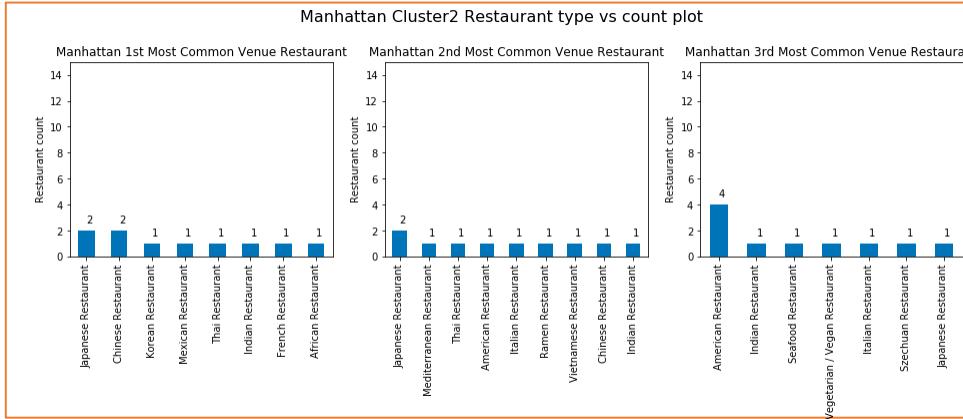
5.2 Analyze Queens all five clusters along with their restaurant category:



Select the cluster that has maximum number of restaurant categories for 3 most common venues:
To make decision simple, we will choose again these two highlighted clusters among Queens five clusters.



6. Best cluster offering a greater number of multi-cultural food restaurants:



From previously selected four clusters, it can be concluded that “Queens cluster2 ” offers wide variety of restaurant categories and number of restaurants compared to all other clusters.

This clustering information will help multicultural food lovers in making decision about the food locations and having all kinds of different ethnic food at one place or within shorter distances.

7. Conclusion:

- “Queens cluster2” offers wide variety of ethnic restaurant categories and number of restaurants.
- Odd distribution of clusters observed for Queens and Manhattan boroughs.
- Difference between Queens and Manhattan total number of neighborhoods is 31.
- May produce biased clustering results because of unbalanced neighborhoods.