

# **Recommending New Cuisine for a Restaurant**

**Tito Francis**

## **1. Introduction**

### **1.1 Background**

New York City (NYC), often called The City or simply New York (NY), is the most populous city in the United States. With an estimated 2019 population of 8,336,817 distributed over about 302.6 square miles (784 km<sup>2</sup>), New York is also the most densely populated major city in the United States.[6] Located at the southern tip of the U.S. state of New York, the city is the center of the New York metropolitan area, the largest metropolitan area in the world by urban landmass.[10] With almost 20 million people in its metropolitan statistical area and approximately 23 million in its combined statistical area, it is one of the world's most populous megacities. New York City has been described as the cultural, financial, and media capital of the world, significantly influencing commerce,[11] entertainment, research, technology, education, politics, tourism, art, fashion, and sports. Home to the headquarters of the United Nations,[12] New York is an important center for international diplomacy.

It's hard to find the right cuisine for the new restaurant if you want to open new or add a new cuisine to your restaurant menu. It can be influenced by geographical area, popular foods types and number of restaurants serving them. There are many businesses shutting down because they couldn't find the right item(cuisines) that can pique the interest of the customers. If you couldn't find the right cuisine for your restaurant it can cause to shutting down your business. So, it's always important to choose the right cuisine for your restaurant so, you can make profit.

### **1.2 Problem**

The cuisine of New York City comprises many cuisines belonging to various ethnic groups that have entered the United States through the city. Almost all ethnic cuisines are well represented in New York City, both within and outside the various ethnic neighborhoods.[1] New York City was also the founding city of New York Restaurant Week which has spread around the world due to the discounted prices that such a deal offers.[2] In New York City there are over 12,000 bodegas, delis and groceries and many among them are open 24 hours a day, 7 days a week.

If you are a restaurant owner or company wouldn't it be great if you could find out popular cuisine in New York or the least number of cuisines available in a neighborhood, so you can decide on the cuisines you want in your restaurant?

### **1.3 Interest**

The aim of this report is to study and analyze the neighborhoods of NYC and group them into similar clusters and, to analyze those clusters to gather meaningful information. That information can be used to find out popular cuisine in New York or the least number of cuisines available in a neighborhood.

Obviously, Restaurants owners should be interested in this project. This project can recommend a cuisine to be added to existed restaurants or new restaurants to be opened.

## **2. Data acquisition and cleaning**

### **2.1 Data sources**

Data sources below are used for the analysis.

a) a dataset that contains the boroughs and the neighborhoods that exist in each borough as well as the latitude and longitude coordinates of each neighborhood [https://cocl.us/new\\_york\\_dataset](https://cocl.us/new_york_dataset)

b) Utilizing Foursquare API to explore the neighborhoods venues that contain cuisine of restaurant

Data used from Foursquare API

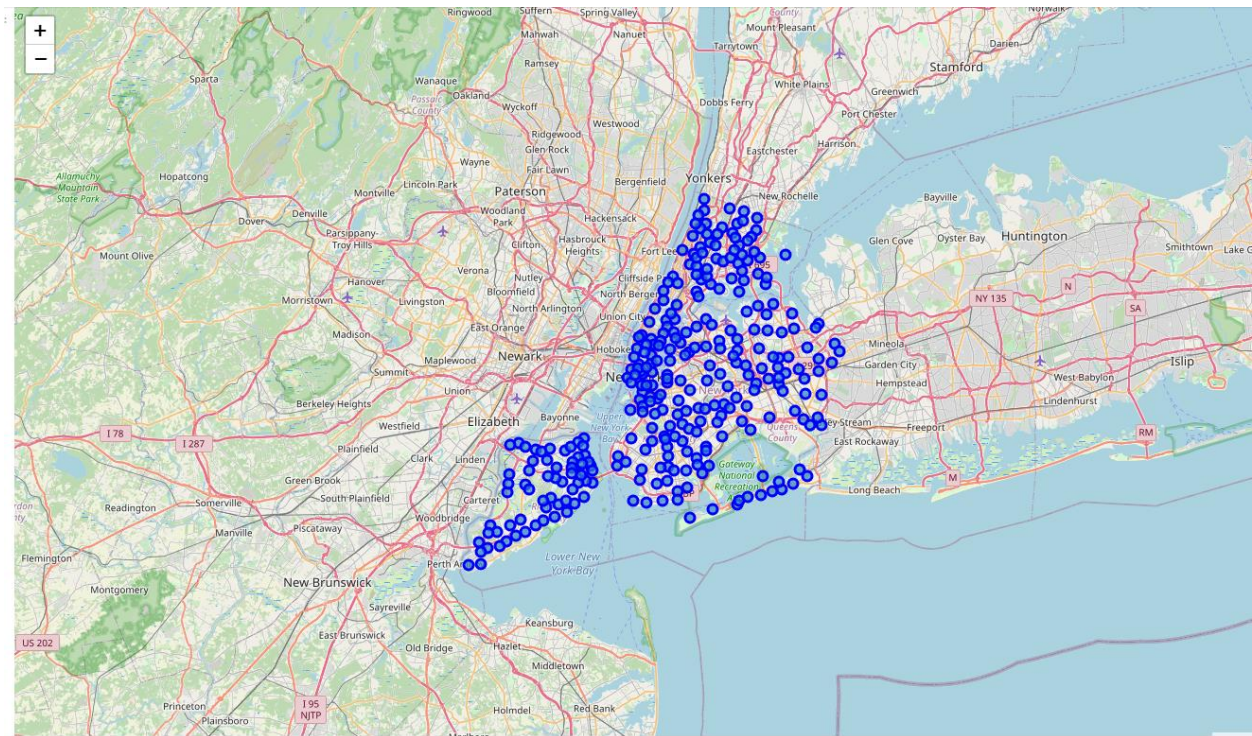
## 2.2 Data cleaning

Data is downloaded from New York dataset and into a table. Then venues from NYC is made into data frame with brought. Then both are combined to form a single table where brought are equal. Then the table is encoded using one-shot coding to produce a table with neighborhoods and its venues. This table is then for this project.

## 2.3 Feature selection

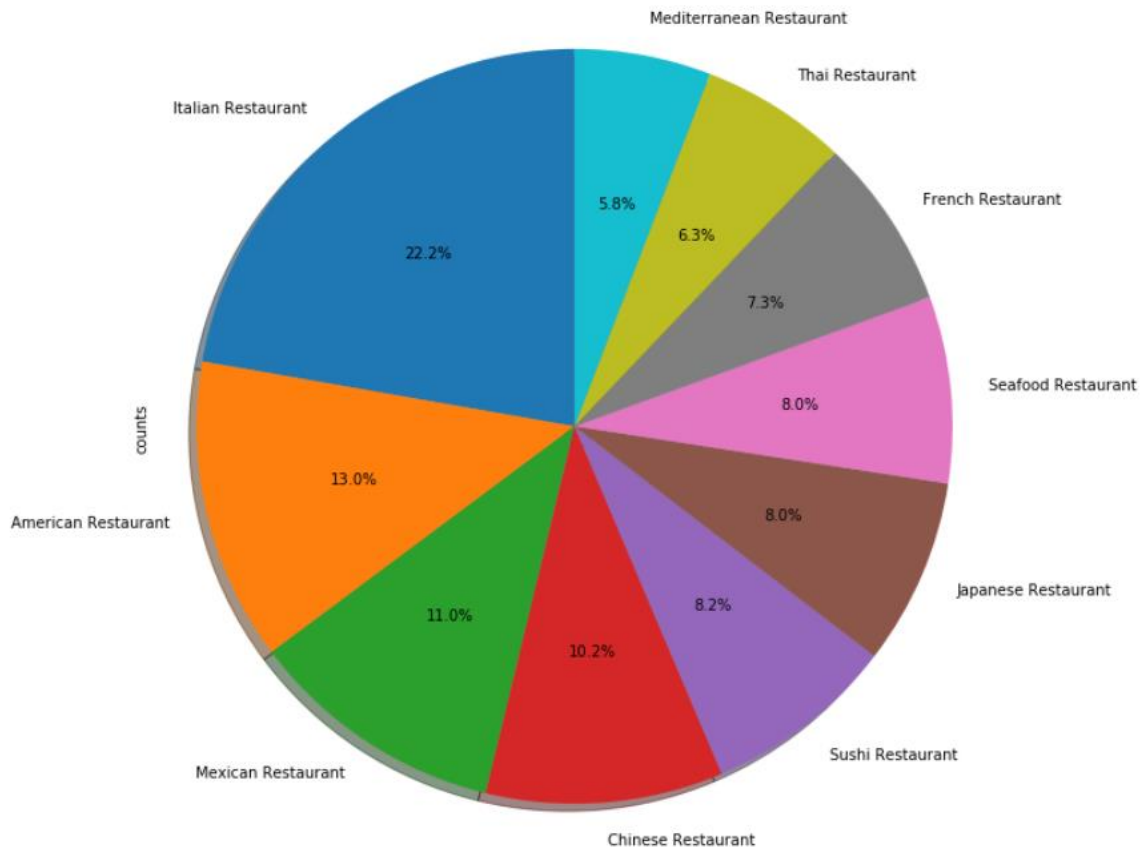
After data cleaning, there were 3019 samples in the data. Upon examining the meaning of each feature(venues), it was clear that there was some venues in the features that can be dropped. Since we are comparing restaurants other venues from the table can be removed. For example, Accessories Store, Antique Shop doesn't affect the cuisines in a neighbourhood. After discarding unnecessary features, I selected venues contain restaurants. It also selected venue type as restaurants from table. This feature doesn't help in deciding the cuisines so, this feature was also dropped. After all, 74 features were selected.

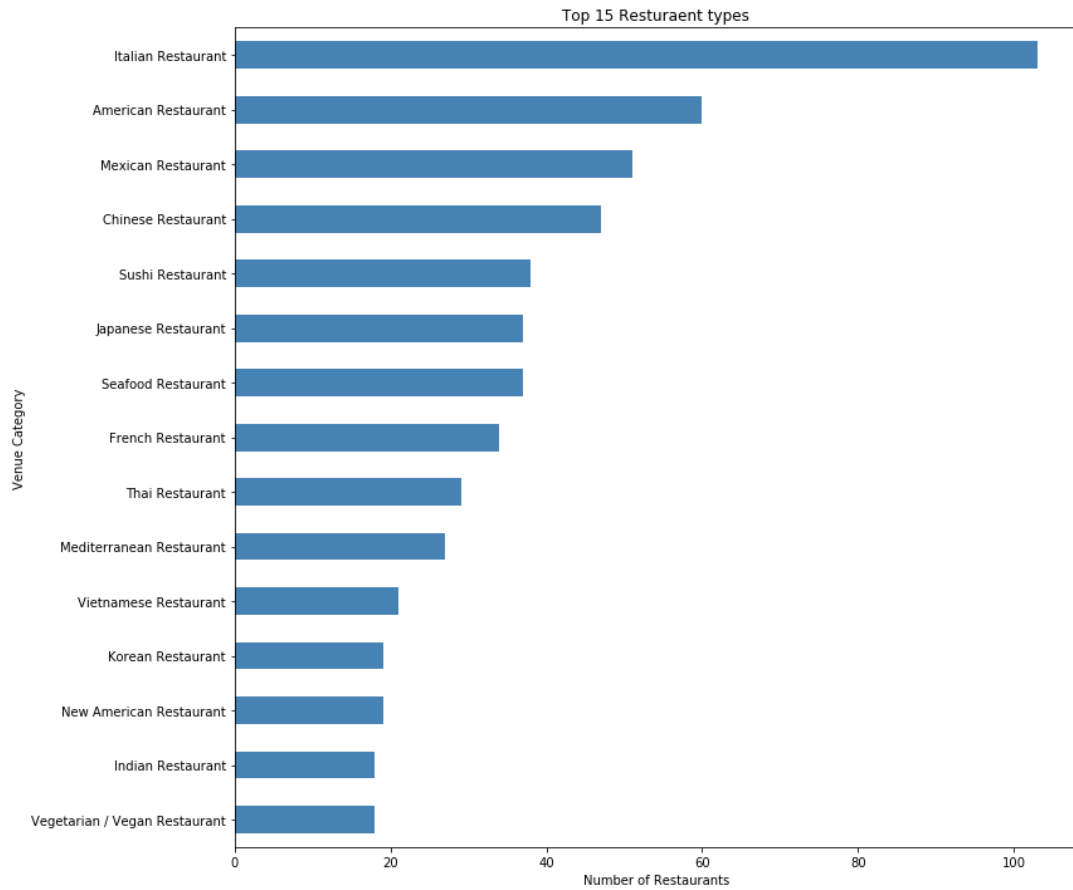
## 3. Exploratory Data Analysis

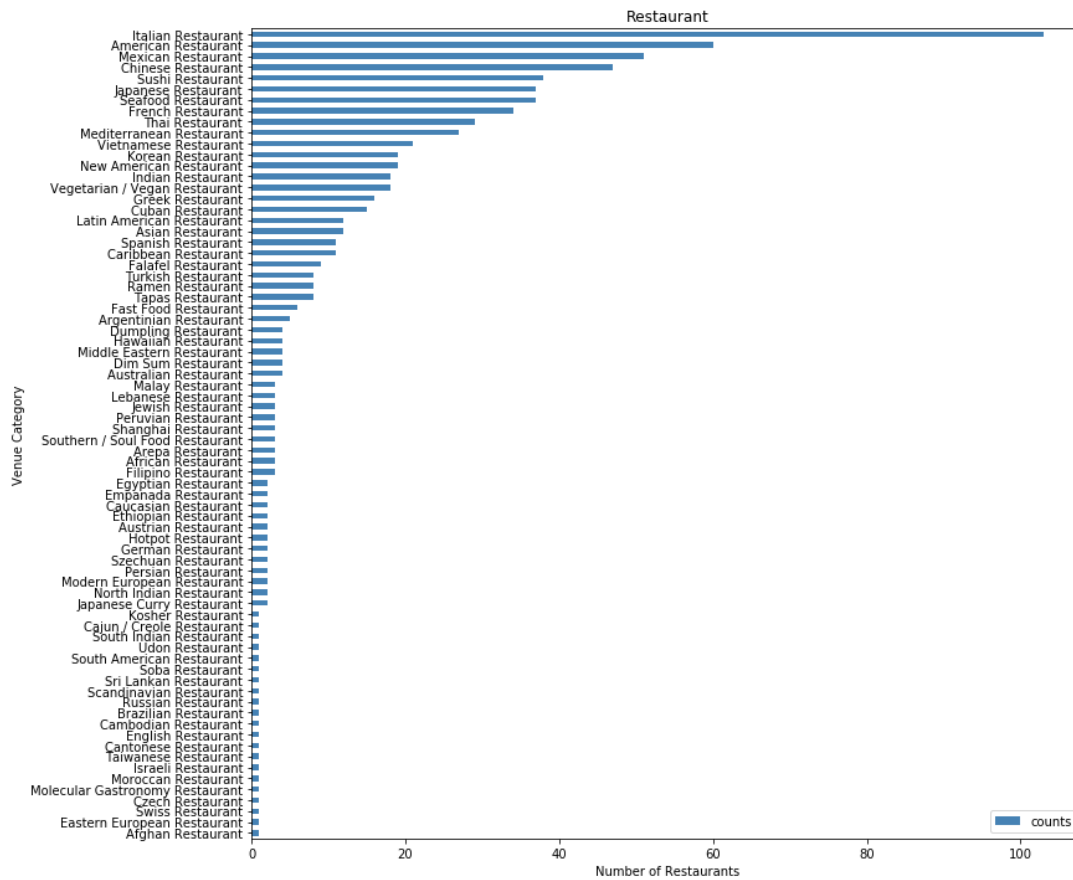


### 3.1 Calculation of Total restaurants types

I choose to calculate the total restaurants. This gives an overall overview of the restaurants in NYC. As we can see it Italian Restaurant is the most popular restaurant. It contributes 22.2% of total top 15 restaurants. The second popular one is American Restaurant with 13%. Coming third is Mexican Restaurant with 11%

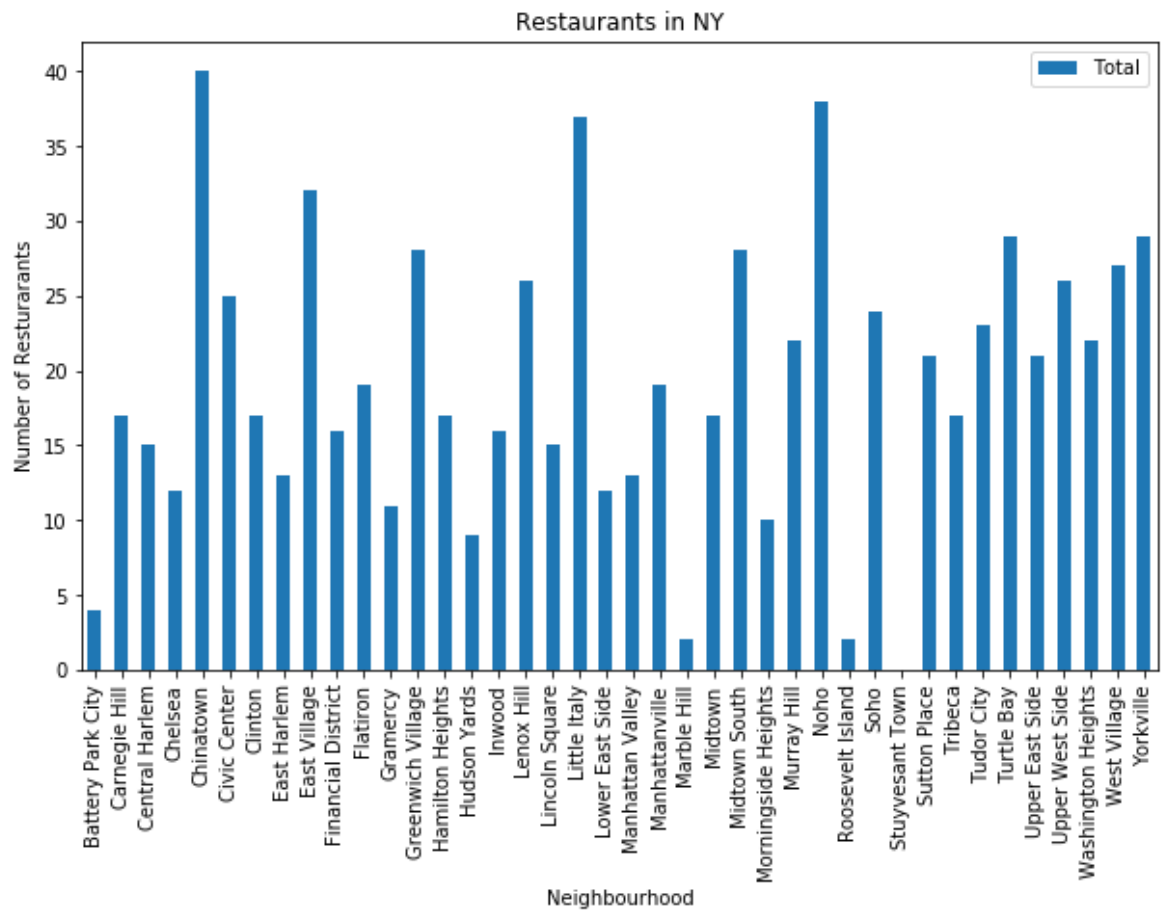


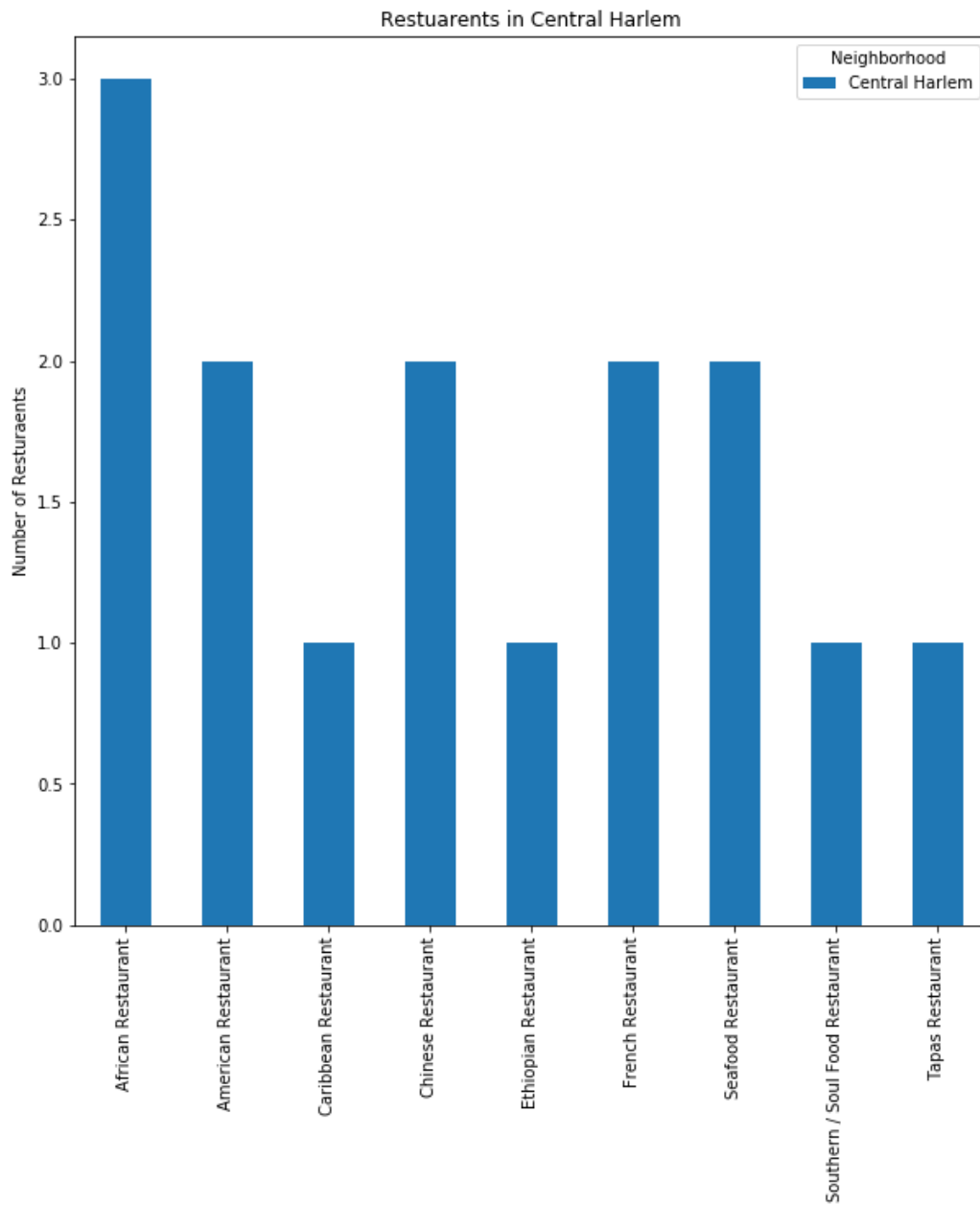


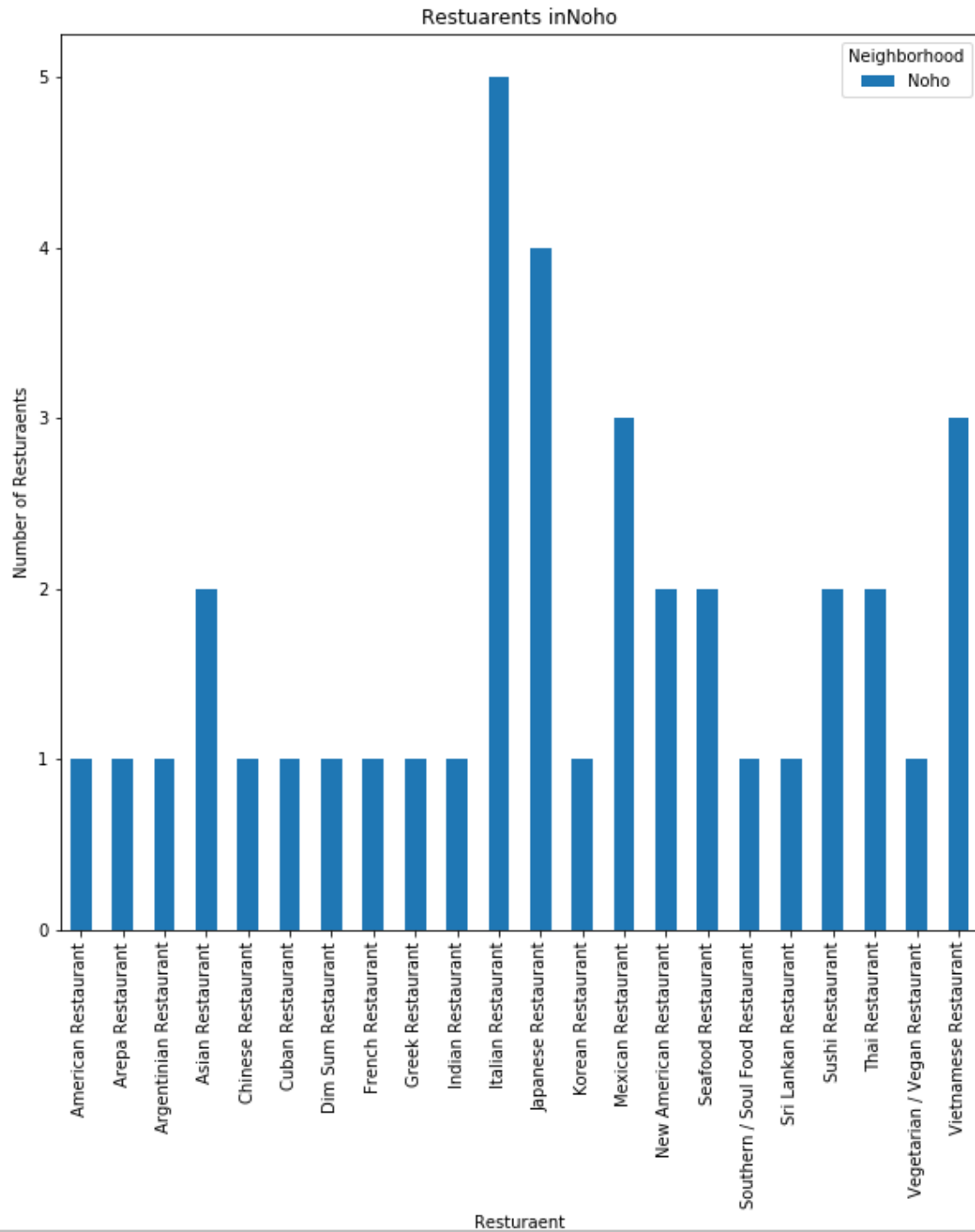


### 3.2 Calculation of restaurant types based on neighborhoods

Here the total number of restaurants in each neighborhood is calculated. We get that the concentration of restaurant in Chinatown is highest. Second highest neighborhoods is Noho. Third highest is Lower East Side







## 4. Clustering

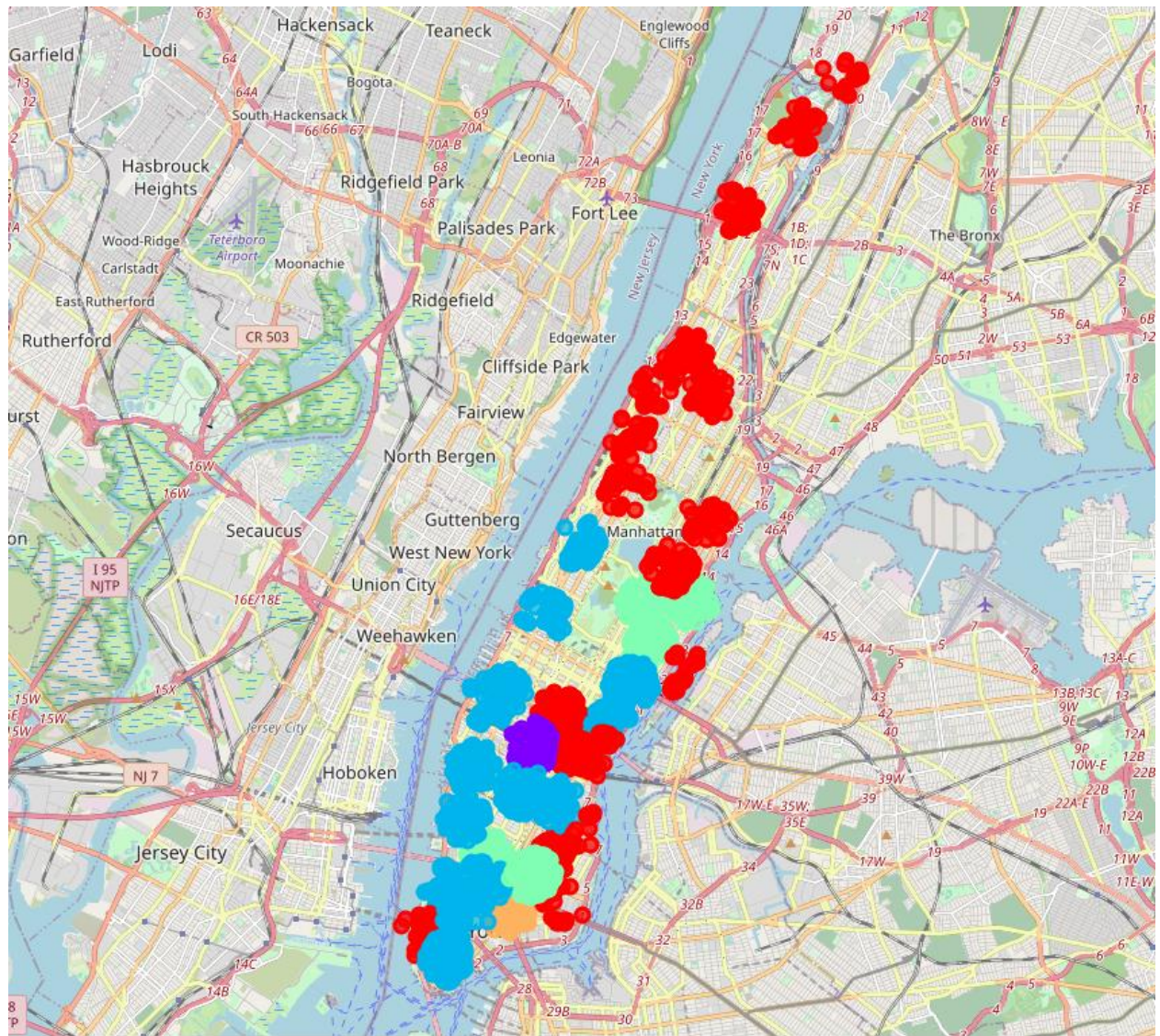
The table of data of neighborhood and restaurants are classified using k means Algorithm

### 4.1 K means

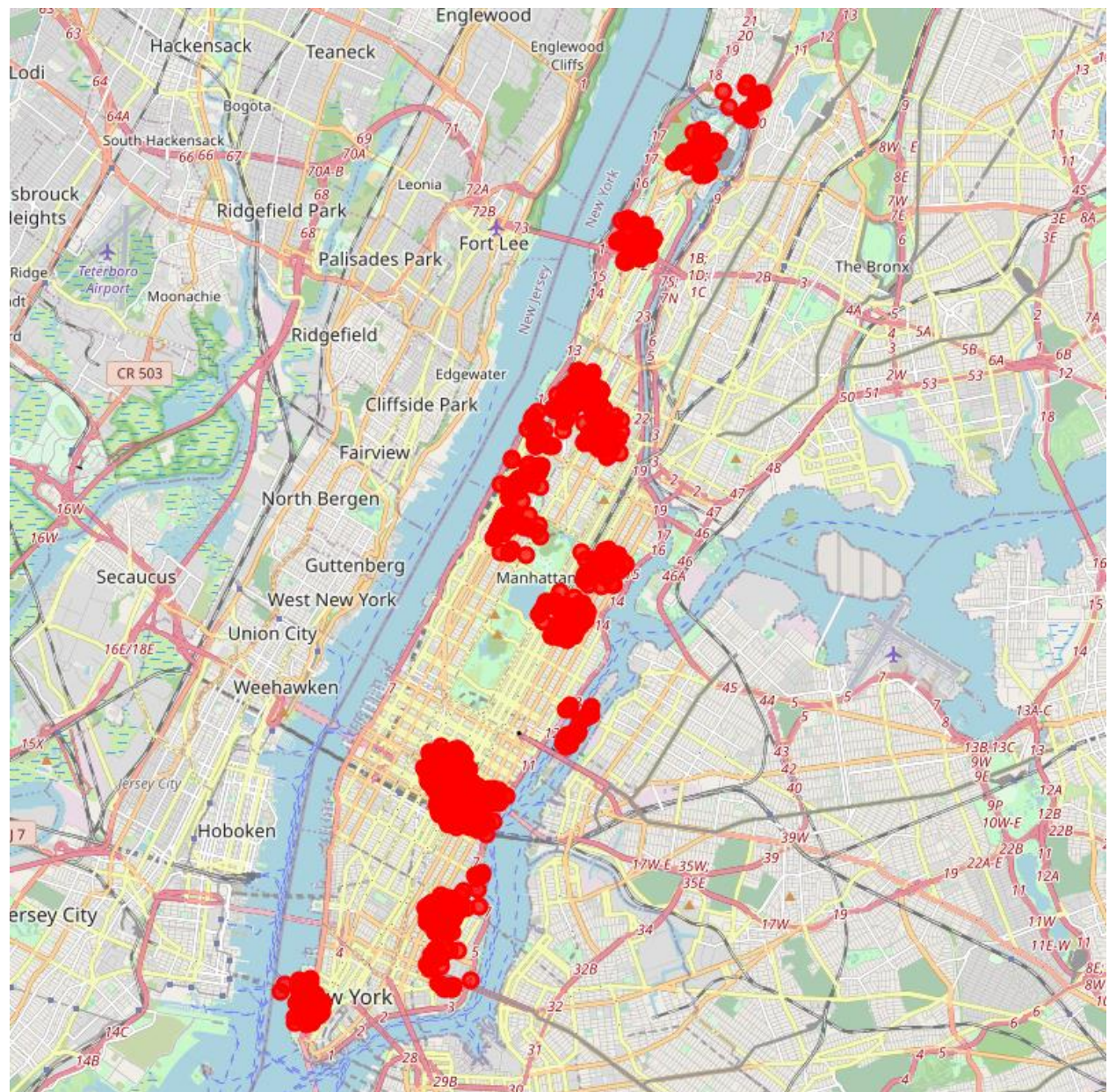


We have some common venue categories in the neighborhoods. We use the unsupervised learning K-means algorithm to cluster the neighborhoods. K-Means algorithm is one of the most common method for clustering in unsupervised learning.

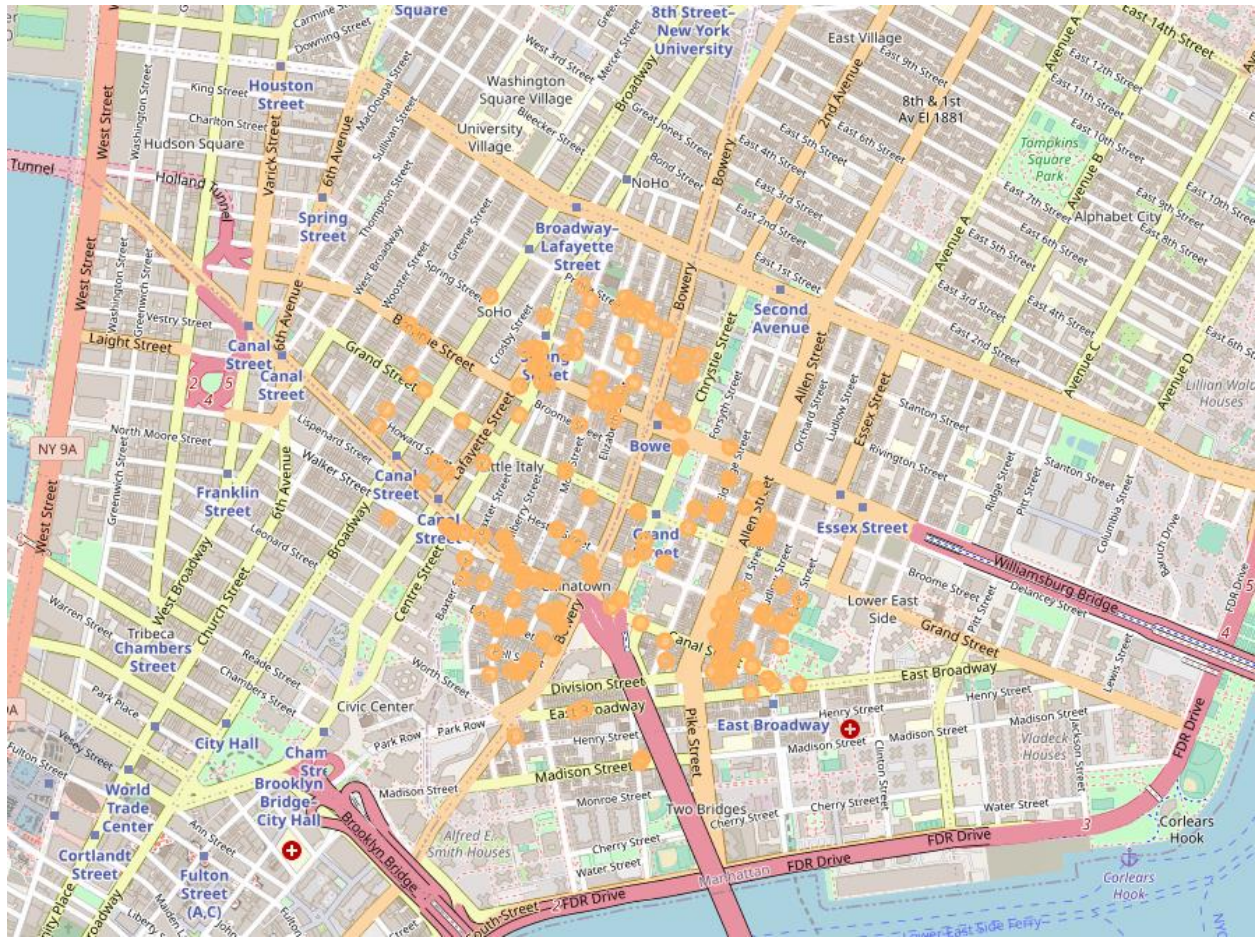
We use a k-clustervalue of 5 to split the neighborhoods into 11 different clusters based on the similarity they have concerning the venues they contain.











## 5. Conclusions

In this study, I analyzed the restaurants and neighborhoods and come to conclusion that Italian restaurant is the most common restaurant in all neighborhood. We also find out the 10 most common restaurant cuisines in all neighborhood. So from this data we can predict the new cuisines to be add to a restaurant menu.

## 6. Future directions

By adding restaurant rating from place can also help in determining the good restaurant type. Also, availability of food product can refine the profit even more.