Coursera Capstone Project

Report Title

Finding an ideal location in Newyork to start a Fine Dining Restaurant

Created by:
Fawaz Mohamed

Problem Description:

The aim of this project is to help investors find a suitable location in New York to start a fine dining Restaurant. The location must fulfill the below 03 criteria:

1. The neighborhood and location should be in one of the 03 most highly populated boroughs in the City. The restaurant will focus on tourists & floating population to nearby attractions.
2. The investors plan to make this restaurant the most popular dining venue in the locality. So, the neighbourhood must not have another established restaurant which is already a top popular venue.
3. The neighborhood of the location must have at least one popular tourist spot or a destination which attracts a lot of traffic.

Target Audience:

One of the main success factors for restaurants is summarized by the age old saying 'Location, Location, Location'.

In fact, having the right location is so important, as it has the power to make or break your business. Restaurant owners and investors always finds it hard to decide on a proper location before making a significant investment.

While traditionally, investors spend money on sending location scouts to do market research about the location, it is time-consuming and expensive . With available data from Four Square and implementing ML algorithms, we can find the most suitable locations, in which they can further do studies on the ground. This will decrease the money and efforts spent in learning many other neighborhoods which are not suitable in the first place.

The results will be used by a set of investors who are looking to start a restaurant in New York. The final results will give them an idea to decide on the ideal location to start the restaurant.

Data Collection

The Data from 03 main sources will be used for this particular project:

1. A dataset giving the list of boroughs and neighborhoods in New York city and their Geographical Co-ordinates. The dataset is available in the link: https://geo.nyu.edu/catalog/nyu_2451_34572
2. A dataset that lists the population of each borough in New York City. New_York_City_Population_By_Neighborhood_Tabulation_Areas.csv will be used for this purpose. The file is available in the link: https://data.cityofnewyork.us/api/views/swpk-hqdp/rows.csv
3. Foursquare API to get the venues and locations in each Neighborhood.

Methodology:

The project followed the steps as mentioned below to find 03 suitable spots for opening a fine-dining restaurant in New York:

1. The latest population figures of New York was downloaded and 03 borough with the highest population where selected.
2. The selected boroughs were analyzed using k-means algorithm and clusters were created, with each cluster having neighbourhoods in a particular borough.
3. The Clusters were first analyzed based on two factors:
    a. Restaurants already available in the cluster.
    b. Public places like beaches, parks, museums etc. available in the clusters.
4. After analysis, one cluster is selected from each borough.
5. A neighbourhood is identified from the cluster, based on the factor that the neighbourhood do not have an already popular restaurant.
6. 03 similar neighbourhoods were identified as suitable location for opening a restaurant that focus on tourists and floating resident population.

Detailed Methodology

Step 01

The population of New York is taken and the 03 boroughs with highest population is identified:

Borough	Year	FIPS County Code	Population
Bronx	76380	190	1385108
Brooklyn	102510	2397	2504700
Manhattan	58290	1769	1585873
Queens	116580	4698	2230722
Staten Island	38190	1615	468730

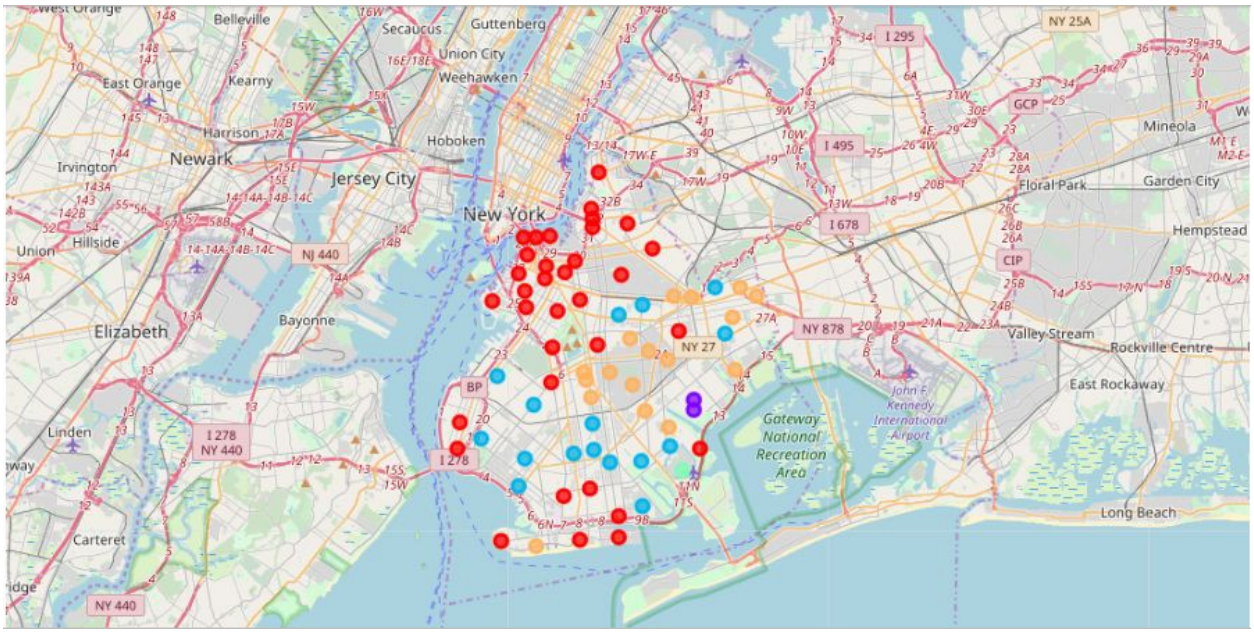
From the above table, Brooklyn, Manhattan and Queens were found to be the borough with the highest population in New York.

So, Brooklyn, Manhattan and Queens were selected for further analysis.

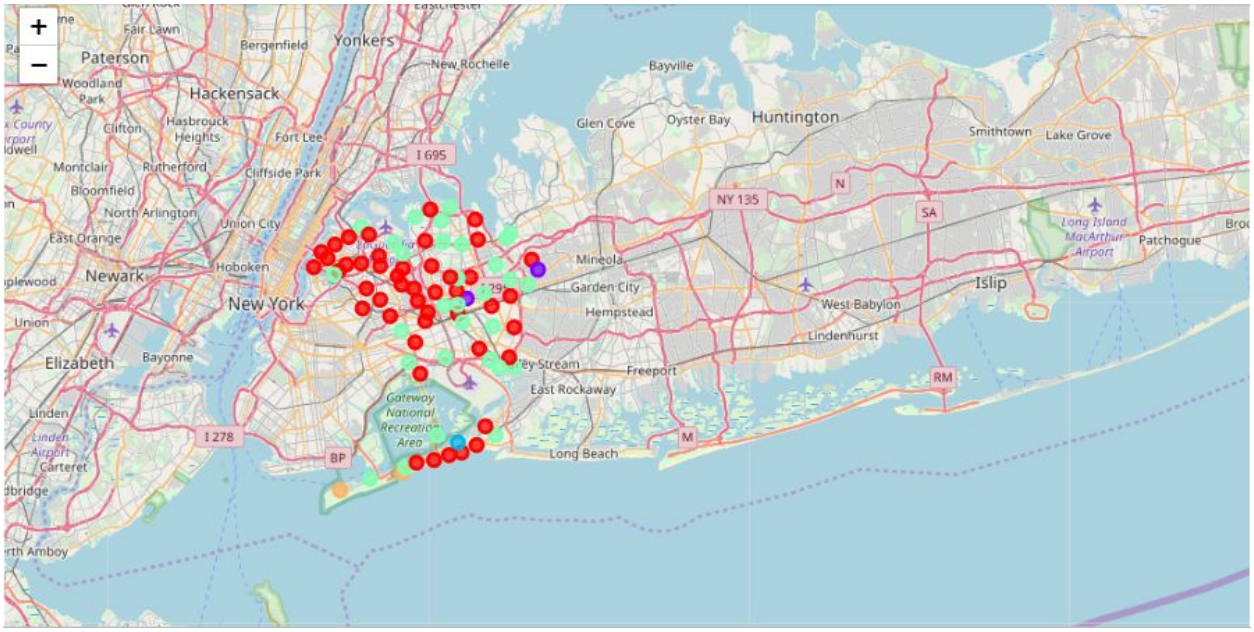
Step 02

Each boroughs were clustered using k-means algorithm as below:

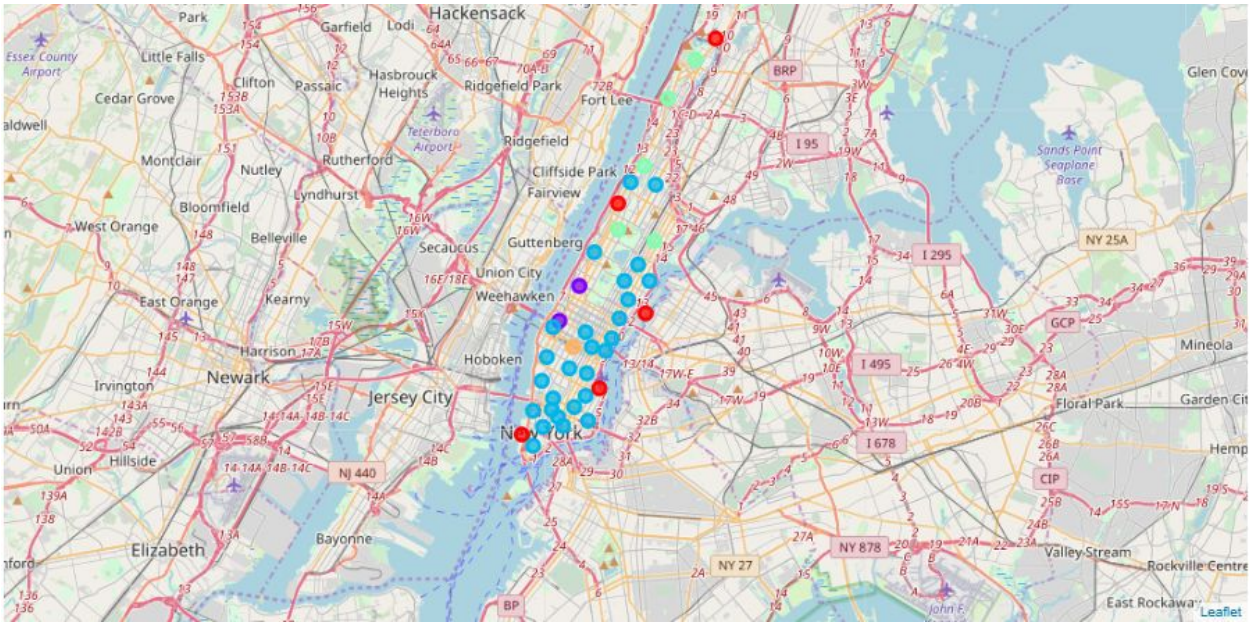
Clusters in Brooklyn



Clusters in Queens



Clusters in Manhattan



Step 03

Each cluster is analyzed based on the most common and popular venues and one cluster from each borough is selected as below:

Brooklyn	
Cluster 01	This is quite a big cluster with neighborhoods having a large number of public attractions including beaches, art galleries, performance venue etc. The cluster also already have a huge number of restaurants. So, finding a neighborhood with less competition is vital for this cluster.
Cluster 02	This is a smaller cluster and do not have any public attractions to attract traffic to the area. So, according to our established criteria, this cluster is not suitable.
Cluster 03	This cluster already have a huge number of restaurants and few public places for attracting potential customers. So, this cluster is not suitable.
Cluster 04	Smallest cluster having a pool and a lake. Since, it's the smallest cluster, will not be considered for future studies.

Cluster 05	This cluster have theme park as an attraction. However, there is a lot of restaurants specialized in various cuisines, giving a high chance of competition for a restaurant.

Queens	
Cluster 01	This is the biggest cluster in this borough with a lot of neighborhoods. But the cluster do not have any major tourist attractions making it unfit for further consideration.
Cluster 02	This cluster has only two neighborhoods, no tourist attractions and will not be good for further considerations.
Cluster 03	This cluster has one neighborhood, and only a park as a major public attraction.
Cluster 04	This cluster already has a lot of restaurants as popular destinations. Do to the possibility of competition, this cluster will not be considered.
Cluster 05	A smaller cluster, but two neighbourhood has beaches which might attract tourists and floating traffic. Moreover, there are no other restaurants being identified as a major venue, as such this cluster can be considered further.

Manhattan	
Cluster 01	This cluster neighborhoods have parks as public places and do not have too many restaurants identified as popular places.
Cluster 02	This cluster have theater, performance arts venue, library, Opera house, Concert halls etc. and would be neighborhoods that attract tourists and floating populations. This could be a cluster suitable for further analysis.
Cluster 03	The neighborhoods in this cluster have a lot of various restaurants making it a venue with higher competition to establish a restaurant
Cluster 04	Already has a lot of cafes, restaurants and coffee shops. Not suitable for consideration.

Cluster 05	Has only one neighborhood and a Korean restaurant is already a popular venue of the neighborhood. So cannot be considered.

Step 04:

Based on the factors mentioned above, Cluster No. 1 in Booklyn, Cluster No. 05 in Queens and Cluster No. 02 in Manhattan are shortlisted as the Clusters for further study.

Step 05:

All neighborhood in the shortlisted Clusters which already has a restaurant as the most popular venue is black listed and the remaining is considered further. After analysis, the below neighborhoods are selected based on the factors mentioned below:

Borough	**Neighborhood**	**Reason**
Brooklyn	Starrett City	The Neighborhood has a shopping mall, a supermarket and a river identified as the most common locations. Doesn't have any restaurants in the most popular locations category.
Queens	Breezy Point	Has a beach, a monument/landmark that attract traffic and also has an event space. Do not have any established restaurant identified as a popular destination.
Manhattan	Lincoln Square	Have a lot of attraction to generate tourist traffic and floating population, including Theaters, plaza, concert halls, performance art theatres etc.

Results:

Based on K-mean clustering, the following three neighborhoods are identified as suitable neighborhoods for opening a fine dining restaurant that attracts tourists:

Borough	Neighborhood
Brooklyn	Starrett City
Queens	Breezy Point
Manhattan	Lincoln Square

Discussion:

Though the above locations are selected as ideal locations for opening a restaurant, it has to be noted that the locations are identified based on the presence/absence of other restaurants and attractions in the locality. Many other relevant factors including rent in the area, local authorities approval/restrictions, space availability, etc, were not considered for the purpose of this study.

Conclusion:

This study considered various boroughs in New York and identified Starrett City, Breezy Point and Lincoln Square as ideal locations for starting a fine-dining restaurant in the City.