**Appropriate Neighborhood for a Restaurant**

Yasser Khalil

31-Jul-2020

**1. Introduction**

**1.1 Background**

When opening a project, there are a lot of risks where this project may fail. One of the risks is that rate of supplier is already higher than the demand so opening such project will not gain the expected profits.  
applying this idea to restaurants, when opening a restaurant in an area which already have many restaurants many of them have their reputation would be a great challenge. Therefore, opening in a neighborhood where there are just a few restaurants so the demand is higher than the supply would be more feasible

**1.2 Problem**

The problem is that this point may be missing from the mind. We can see much restaurant opening in a street full of well-know restaurants and no one would ever notice it.

**1.3 Interest**

Project Managers would need such info to notice what are the most common venues in this neighborhood and to check if opening their project would be feasible and succeed or not.

**2. Data acquisition and cleaning**

**2.1 Data sources**

We used a dataset from Coursera’s IBM data science course to get New York city neighborhoods along with the coordinates for each neighborhood and we used the Foursquare API to explore the neighborhoods and get information on the neighborhood venues

**2.2 Data Cleaning**

The data had all New York city neighborhoods, for simplifying purposes we minimized the data to Manhattan island only and filtered all other brogues.

**3. Exploratory Data Analysis**

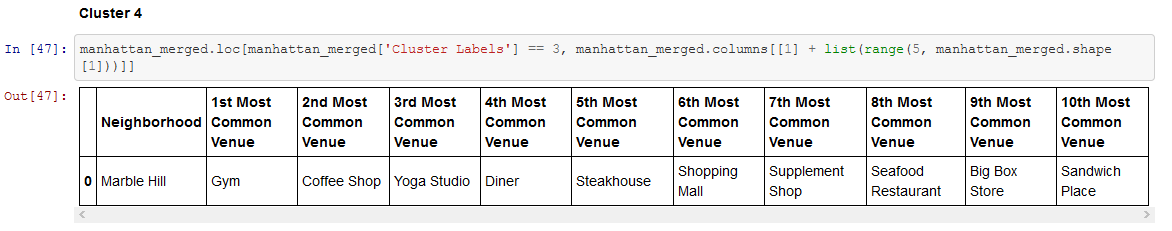
Whereas this report is devoted for the restaurant. The way the project works is by clustering all the neighborhoods and their venues without aiming to any target. Which means that the clusters produced could be used for many other purposes like checking where to open a hospital, where should the fire trucks be stationed, etc.…  
therefore the technique taken for this approach would be clustering and we used K-Means algorithm.

**4. Results**

The relation between the clusters and the action that will be taken is based upon the most common venues. As we can notice some clusters like [1,2] have restaurants & Bars as the Top 3 Common venues. Which would be a disadvantage to open a restaurant in that neighborhood.



Whereas other clusters like Cluster 4 has Gyms as the most common Venues and restaurants are not listed as the top 3 commons.

  
therefore, opening a restaurant in that district [maybe specialized in healthy food] would be a good idea.

Another approach I would take is by using the clusters to analyze the venues by key words and check whether the top venues contains anything related for the desired search.



**5.Conclusions**

In this study, we could get Manhattan neighborhoods and cluster them based on their venues and the clusters where clearly shown on the map using Folium API. This clustering would help us to decide if we should open a project/shop in some neighborhood based on the customer need. And we can use those clusters to generate many other actions to be taken.

**6. Future directions**

An idea to work on in the future is analyzing the whole venues and giving weight to each order descendible way. Ex: the first common venues gets the most weight and the less takes less weight. Afterwards we can check the weight of each service in a specific neighborhood and decide on the action based on those weight ([where restaurants, Bars, Cafes & Diners can be all specified into one service [food & beverages])