Il Carnicero (steakhouse) from Milan - Ibiza and now New York

Background and Business Problem Statement

I was having dinner at my favorite Steakhouse in Milan (El Carnicero) and i met the owner of that restaurant. We started to talk and he told me that he already opened another restaurant 2 years ago in Ibiza and now he wants to expand also in the United States of America - New York. He is very curious about the steakhouse market in New York and he is very curious what is the best neighbourhood to open his new restaurant. As he knows that he has a very good restaurant he is searching to open near other stakehouses as he knows that his receipes are highly competitive.

His question is:

"What neighborhood in NYC would be the best fit for opening up a steaknouse shop based on the places already freqvented by other "meat lovers".

Data Sources for the Project:

I will use the following datasets that are already used in the course.

1. NY Neighborhood Data -
   https://ibm.box.com/shared/static/fbpwbovar7lf8p5sgddm06cgipa2rxpe.json &
   https://geo.nyu.edu/catalog/nyu_2451_34572
2. Foursquare Location Data (Venues [Restaurants, Community Centers], TOP tips, Favorites, User Experience, etc) will be used to cluster, segment, target, and position to craft recommendations for suitable location to open up a vegan ice-cream shop. Using the data available in the above sources, I will be conducting clustering and neighborhood based analysis leveraging primarily Foursquare APIs and tools such as KNN and relevant unsupervised learning methods to deliver recommendation options to the client.

Here is some relevant examples of tailored data that will benefit the the client on making the decision which place would be higher populated with his targeted audience:

a) There are many vegan eating places in NY with the top rating being 9.3 for abcV Restaurant.
b) There are also tens of neighborhoods truly deliver pristine eating and relaxing environment for meat lovers, which makes them highly attractive for the target audience of the client.
c) Leveraging the recommender systems capabilities enabled by Foursquare to conduct compare / contrast the data points to derive the recommendations.

Here are the specific steps that will be deployed to work with the data derived from Foursquare and the approach will cover NYC and also the category of "Vegan Community Centers"

1. Search for a specific venue = "Stakehouse"
2. Perform the GET Request and examine the results for Stakehouse eating places
3. Get relevant part of JSON and transform it into a pandas dataframe & review the data for patterns and obvious insights
4. Explore Venues (=Stakehouse related) with Rating of 8.5 (out of 10) or more
5. Get the number of tips for each Venue and focus on what could be a source of valuable information for the Data Scientists (new-hires)
6. Gather frequent user feedback and comments
7. Narrow down the venues and the associated neighborhoods for each city that can form the set of recommendations

Methodology:

The methodology employed is aligned with the Data Science 10 steps program that was discussed during an earlier module of thecourse.

- Obtain relevant vegan places data from Foursquare and clean it for data understanding and grouping etc.
- Explore the data for clusters & patterns of Neighborhoods in Boroughs
- Group the places of high interest into relevant neighborhood and Borough pairings
- Generate the markers on both NYC map to highlight the neighborhoods that are in focus.
- Use Foursquare venue, Categories etc to enable the profiling the primary & focus neighborhoods.
- Generate the short-listed neighborhoods for NYC.

Results:

Here is the final short-list of neighborhoods with that accomodate the most for vegans in NYC and there fore have higher presence of targeted audience of the client:

```
•   Bushwick and Williamsburg        2
•   Chelsea and Clinton              8
•   Gramercy Park and Murray Hill    3
•   Greenpoint                       2
•   Greenwich Village and Soho       2
•   Lower East Side                  5
•   Northwest Brooklyn               2
•   Upper East Side                  1
```

It was clear that Greenwich, Greenpoint, Chelsea and Clinton and Lower East Side were highly populated with stakehouse in NYC.

Based on my observation, the 3 most recommended neighborhood to open up the ice-cream shop that fits our client's search criteria are Greenwich, Greenpoint and Lower East Side with the true winner Greenwhich.
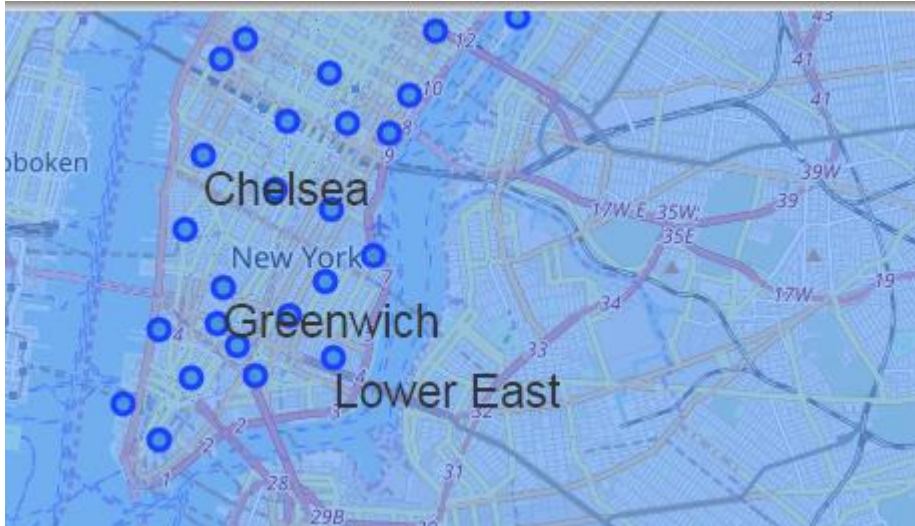
While this project was able to derive specific recommendations in terms of places in NY that are highly populated with vegans, it has limitations given that the assumptions for the model are simplistic in nature. The main variable of focus was only Food & Leisure where as vegans may be interested in a number of other amenities and interests that will make it a multi-variable focused data science analysis. From my perspective the project provided a tremendous opportunity for me to explore and apply most of the knowledge and insights gained during the previous 8 modules of this course.

Summary & Conclusion:

With all of its limitations, this project can be useful for the client to get a quick glance of the neighborhoods in NYC, that would have a higher population of his targeted audience.