Opening a Michelin Guide starred restaurant in New York City

I. Introduction & Business Problem:

A. Background

An investor wants to open a restaurant in New York City (NYC) with the ambition to be rated in Michelin Guide.
New York City is the most populous city in the USA. New York City is home to "nearly one thousand of the finest and most diverse haute cuisine restaurants in the world", according to Michelin. As of 2019, there were 27,043 restaurants in the city, up from 24,865 in 2017. This means that the market is highly competitive. Therefore, the project of opening a new restaurant in New York City needs to be analyzed carefully.

B. Problem

The business decision restaurant project is based on multiple factors. One key factor is the ability to distinguish yourself from the competition.
Knowing that, the US investor asks to a consultant to provide him with the landscape of the restaurant in NYC by type of cuisine.
The success of the project will be a good recommendation of borough/Neighborhood choice to the US investor based on the lack of such restaurants in the recommended area.

C. Interest

This report would be useful to anyone who would be interested to invest in a 'haute cuisine' restaurant in New York City.

II. Data acquisition and cleaning

A. Data sources

We will use the following sets of data in our project:

Data 1
We will retrieve New York City boroughs and neighborhoods names and locations from the following dataset https://geo.nyu.edu/catalog/nyu_2451_34572

	Borough	Neighborhood	Latitude	Longitude
0	Bronx	Wakefield	40.894705	-73.847201
1	Bronx	Co-op City	40.874294	-73.829939
2	Bronx	Eastchester	40.887556	-73.827806
3	Bronx	Fieldston	40.895437	-73.905643
4	Bronx	Riverdale	40.890834	-73.912585

Data 2
We will retrieve from Wikipedia the list of Michelin starred restaurants in New York City:

https://en.wikipedia.org/wiki/List_of_Michelin_starred_restaurants_in_New_York_City

List of Michelin starred restaurants in New York City

From Wikipedia, the free encyclopedia

This article contains a complete list of Michelin starred restaurants in New York City and Westchester County. The 2006 edition was the first edition of the *Michelin Guide to New York City* to be published. It was the first time that Michelin published a *Red Guide* for a region outside Europe.[1] In the 2020 edition, the Guide began to include restaurants outside the city's five boroughs, adding Westchester County restaurants to its listing.[2]

Contents [hide]
1 List
2 See also
3 References

List [edit]

Name	Borough/County	2006 (39)[3]	2007 (39)[4]	2008 (42)[5][6]	2009 (42)[7][8]	2010 (55)[9]	2011 (57)[10]	2012 (62)[11]	2013 (67)[12][13]	2014 (67)[14]	2015 (73)[15]	2016 (76)[16]	2017 (77)[17]	2018 (72)[18]	2019 (76)[19]	2020 (76)[20]
15 East	Manhattan								❁	❁	❁					
A Voce Columbus	Manhattan		❁	❁		❁	❁	❁	❁	❁			Closed[21]			
A Voce Madison	Manhattan						❁	❁	❁	❁				Closed[22]		

Data 3
New York City Michelin starred restaurants data will be utilized as input for the Foursquare API, which will be leveraged to provision restaurant information for each neighborhood.

	Venue	Borough	S_rate	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue Latitude	Venue Longitude	Venue Category
0	Ai Fiori	Manhattan	1	Midtown	40.754691	-73.981669	40.750075	-73.983784	Italian Restaurant
1	Atera	Manhattan	2	Tribeca	40.721522	-74.010683	40.716752	-74.005712	Molecular Gastronomy Restaurant
2	Atomix	Manhattan	2	Murray Hill	40.748303	-73.978332	40.744306	-73.982945	Korean Restaurant
3	Bâtard	Manhattan	1	Tribeca	40.721522	-74.010683	40.719624	-74.005788	Modern European Restaurant
4	Blanca	Brooklyn	2	East Williamsburg	40.708492	-73.938858	40.705033	-73.933774	New American Restaurant

III. Methodology

A. Analytic Approach

To achieve our goal we will:

- Pin point the stared restaurant in NYC,
- Cluster them,
- Analyze the type of cuisine in each neighborhood.

B. Exploratory Data Analysis

1. Data 1 – NYC neighborhoods

We perform the following actions:

- we load the data and explore data from the json file,
- we transform the data into a panda dataframe,
- we use geopy and folium libraries to create a map of NYC with the neighborhoods located
- this data will be merged in our final map.

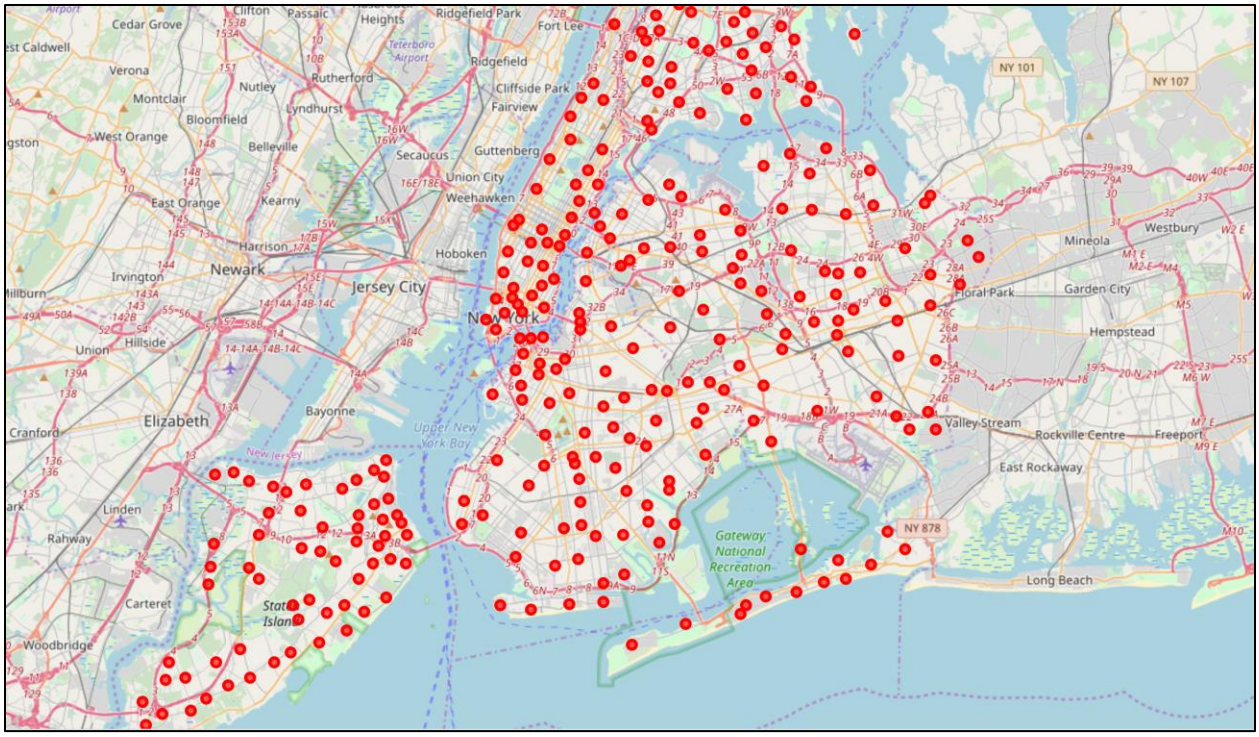


Figure 1 - New York City neighborhood visualization

2. Data 2 - list of Michelin starred restaurants in New York City

We perform the following actions:

- The data has been retrieved using BeautifulSoup.
- We order the columns,
- We dropped the restaurant that lost their stars in 2020
- We formatted the star rating column to keep only the number of stars,
- We will cross this data with the data received from foursquare.

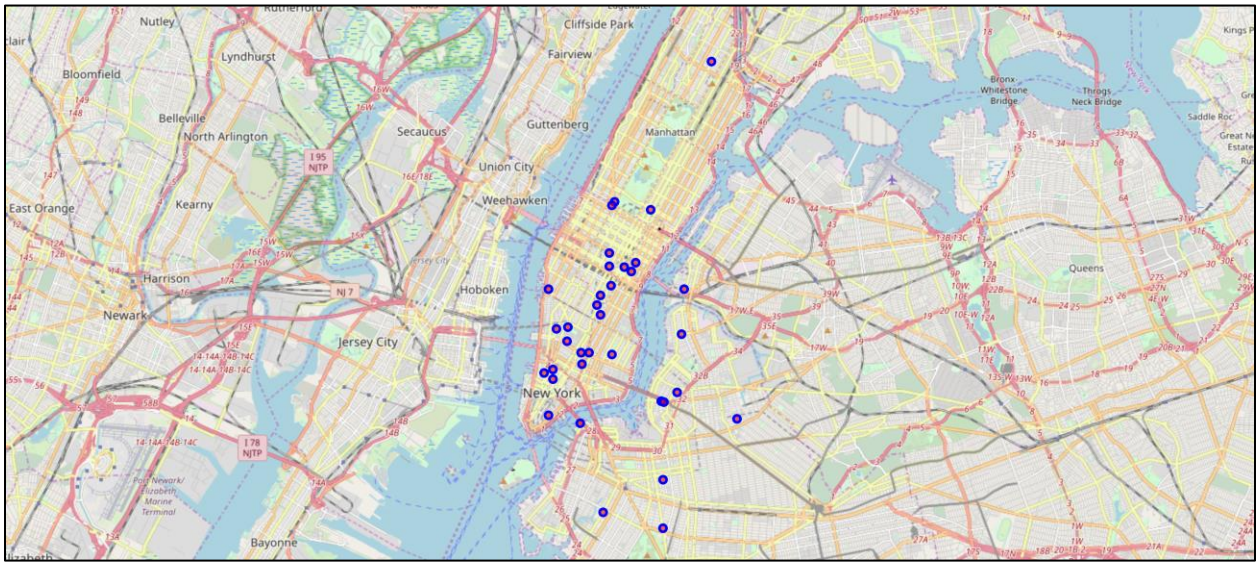
	Venue	Borough	S_rate
0	Agern	Manhattan	1
1	Ai Fiori	Manhattan	1
2	Aldea	Manhattan	1
3	L'Appart	Manhattan	1
4	Aquavit	Manhattan	2

3. Data 3 - Restaurants information using the Foursquare API

We perform the following actions:

I. We retrieved all information on the NYC neighborhood restaurant,
II. We crossed the information with our list of restaurant to retrieve only the stared restaurant including latitude, longitude and type of cuisine of our restaurant

III. We put them on map to have visualization



IV. We clustered them into 3 groups using the K-means method

	Venue	Borough	S_rate	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue Latitude	Venue Longitude	Venue Category	Labels
0	Ai Fiori	Manhattan	1	Midtown	40.754691	-73.981669	40.750075	-73.983784	Italian Restaurant	2
1	Atera	Manhattan	2	Tribeca	40.721522	-74.010683	40.716752	-74.005712	Molecular Gastronomy Restaurant	0
2	Atomix	Manhattan	2	Murray Hill	40.748303	-73.978332	40.744306	-73.982945	Korean Restaurant	2
3	Bâtard	Manhattan	1	Tribeca	40.721522	-74.010683	40.719624	-74.005788	Modern European Restaurant	0
4	Blanca	Brooklyn	2	East Williamsburg	40.708492	-73.938858	40.705033	-73.933774	New American Restaurant	1

V. We calculated the mean to have the center of each cluster

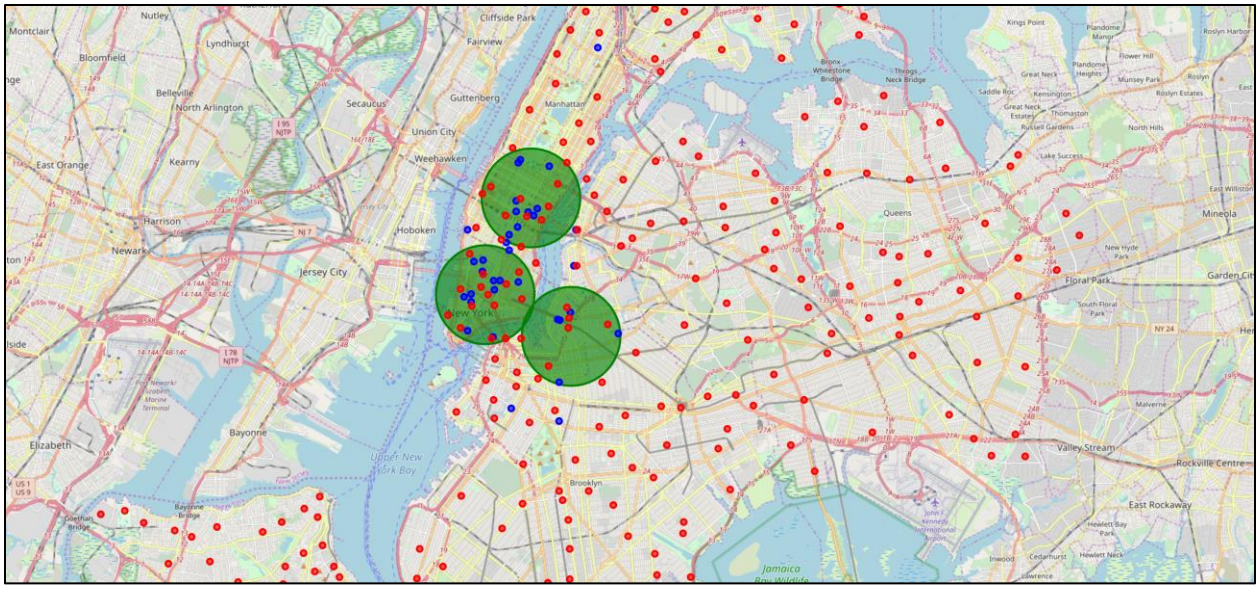
	Labels	Neighborhood Latitude	Neighborhood Longitude	Venue Latitude	Venue Longitude
0	0	40.719634	-74.000474	40.719451	-73.998869
1	1	40.756243	-73.974307	40.755113	-73.976354
2	2	40.704311	-73.957164	40.704032	-73.956808

VI. We used the “one hot encoding” to visualize each type of cuisine in each Neighborhood

	Neighborhood	American Restaurant	Asian Restaurant	French Restaurant	Italian Restaurant	Japanese Restaurant	Korean Restaurant	Mediterranean Restaurant	Mexican Restaurant	Modern European Restaurant	Molecular Gastronomy Restaurant	New American Restaurant	Restaurant	Spanish Restaurant	Steakhouse	Sushi Restaurant	Thai Restaurant	Wine Bar
0	Wakefield	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
1	Co-op City	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
2	Eastchester	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
3	Fieldston	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
4	Riverdale	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0

IV. Results

As a result we obtained the following map:



In red the Neighborhoods, in blue the restaurants and in green the clusters.

V. Discussions

1. We can see that the restaurants are concentrated in the south of Manhattan but still we have restaurants outside of this zone. We recommend investing outside of this area.
2. The analysis of the type of restaurant by Neighborhoods showed that types of restaurant are very diverse and there is no specific type of restaurant more than another. We think that we can open any type of restaurant.

VI. Conclusion

We identified that the density of restaurant was higher in south Manhattan and that there was no specific type of restaurant that was more represented. But the business decision to open a restaurant is not only on the parameter selected, but it will give early insights to the investor to perform more data analyses. We recommend for deeper analyses when the type of restaurant has been selected to investigate the menu the restaurant of this type.