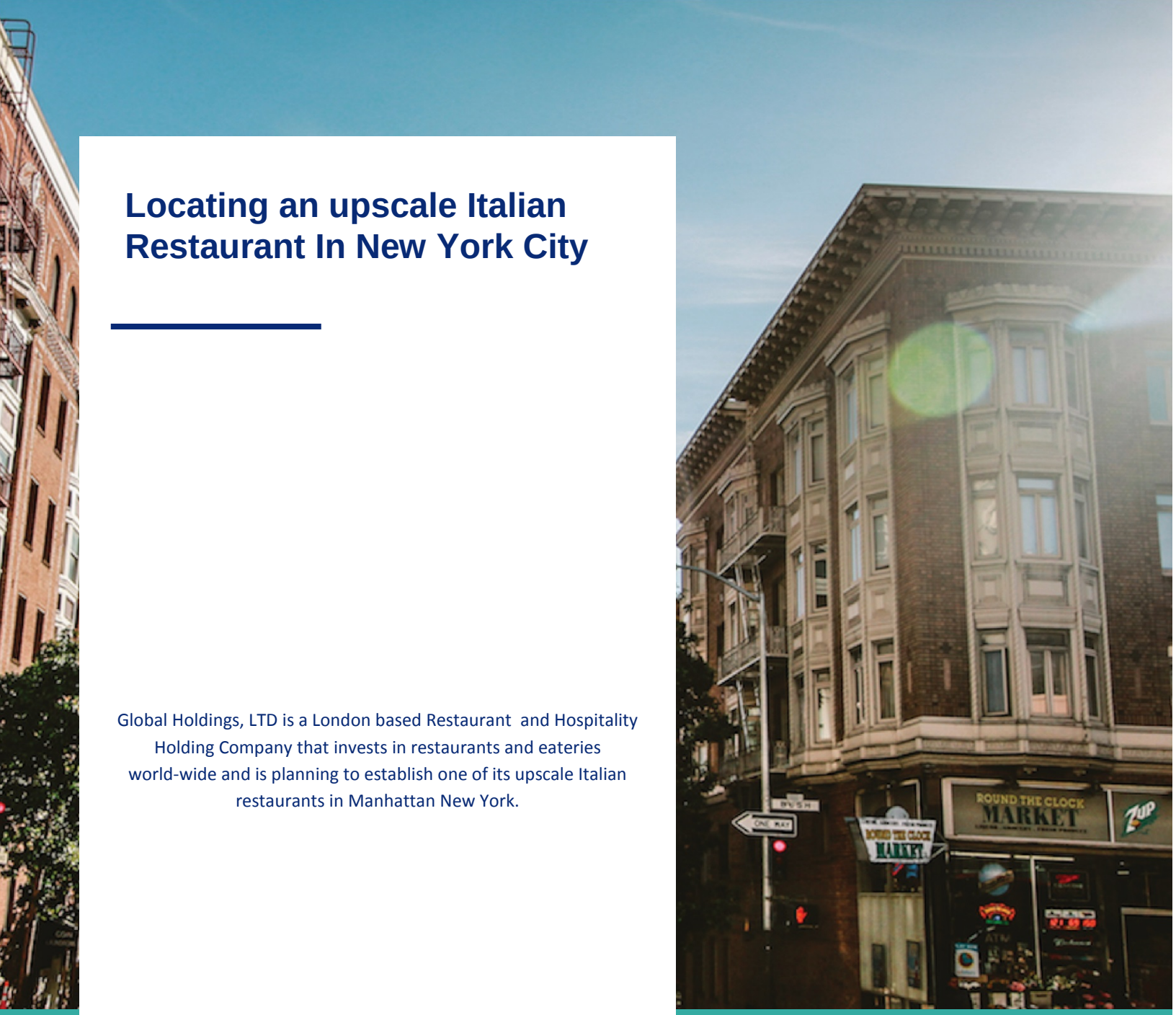
Locating an upscale Italian Restaurant In New York City



Global Holdings, LTD is a London based Restaurant and Hospitality Holding Company that invests in restaurants and eateries world-wide and is planning to establish one of its upscale Italian restaurants in Manhattan New York.

Pages: 14
Word count: 1,909

APRIL 10, 2020

Authored by: Norman Noplock
Global Holdings, LTD
152 - 160 City Road,
London, EC1V 2NX

Executive Summary

Italian food has conquered more nations – or at least their palates – than the Roman Empire ever could have hoped to militarily. Italian restaurants are found all over the world from Tokyo to Toronto. Italian cuisine has skyrocketed in popularity in recent years and approximately 44 percent of Americans eat Italian food 1- 2 times per week.

Opening an Italian restaurant requires extensive research and detailed due diligence including but not limited to, location site analysis and market studies. One must determine the number of Italian restaurants already established in the area under review. Also, what is the predictive analysis that the candidate sites are conducive to a new business. One must also determine if the established competition hinders or enhances the potential of success.

Restaurants are classified into three primary categories: quick-service, midscale and upscale. Quick-service restaurants are also known as fast-food restaurants. Midscale restaurants, as the name implies, occupy the middle ground between quick-service and upscale restaurants. Upscale restaurants offer full table service and do not necessarily promote their meals as offering great value; instead, they focus on the quality of their cuisine and the ambience of their facilities. Fine-dining establishments are at the highest end of the upscale restaurant category and charge the highest prices.

Global Holdings, LTD (the Company) is a London based Global Restaurant and Hospitality Holding Company that invests in restaurants and eateries world-wide. To date the Company has investments is seventeen countries with plans to expand its number of business entities by an annual growth factor of 12%.

Upscale Italian Restaurant

My role as a Business Development executive for the Company, includes identifying, through a detailed due diligence process, locations suitable for new restaurants and other eateries. I have been recently assigned the task of conducting a high level review of locations in New York City that are suitable for an upscale Italian Restaurant.

Locating an upscale Italian Restaurant In New York City

Introduction

Restaurant location research requires extensive and detailed due diligence including but not limited to location site analysis. The following is a summary check list used by the Global Holdings, LTD (the Company) to determine if a site qualifies as a location for one of its a restaurants.

- Anticipated sales volume. How will the location contribute to sales volume?
- Accessibility to potential customers. Consider how easy it will be for customers to get into the business. Consider whether or not nearby businesses will generate foot traffic.
- Neighborhoods. Determine if the neighborhood where the site is located has similar and or competitive established businesses. It must be decided if the established competition hinders or enhances the potential of success.
- Restrictive ordinances. Restrictive ordinances that make an otherwise strong site less than ideal, such as limitations on the hours of the day that trucks can legally load or unload.
- Traffic density. A careful study of foot traffic can determine the approximate sales potential of each pedestrian passing a given location. Two factors are especially important in this analysis: total pedestrian traffic during business hours and the percentage of it that is likely to patronize the business.
- Customer parking facilities. The site should provide convenient, adequate parking as well as easy access for customers.

- Proximity to other businesses. Neighboring businesses may influence business volume, and their presence can work for you or against you.

- History of the site. Research the recent history of each site under consideration. Who were the previous tenants, and why are they no longer there?

- Terms of the lease. Be sure you understand all the details of the lease, because it's possible that an excellent site may have unacceptable leasing terms.

- Future development. Check with the local planning board to see if anything is planned for the future that could affect the business, such as additional buildings nearby or road construction.

The focus of this report is a high-level review of:

- Neighborhoods. Determine if the neighborhood where the site is located has similar and or competitive established businesses. It must be decided if the established competition hinders or enhances the potential of success.

Room for one more

The Company has been successful in the past by launching one of its restaurants in locations where there are already similar type businesses. As such, this report includes a final list of neighborhoods where there are existing Italian restaurants and eateries.

In this high-level review I will make use of a New York data Set from the New York University Spatial Data Repository, location data from FourSquare and demographics from the United States Census. The New York Dataset provides Borough, Neighborhood, and location coordinates for all of New York City. I intend to single out one or perhaps two suitable New York Boroughs (there are 5 to select from) and then merge FourSquare venue data in order to determine New York neighborhoods (there are 306) that have established Italian restaurants. US Census data will be added to provide, population, housing and income demographics.

Manhattan New York

I was issued additional instructions to first research Manhattan, a borough of New York city, for neighborhoods that qualify as a potentially suitable locations for an upscale Italian restaurant.

Methodology

I was able to arrive at potential Manhattan neighborhoods, suitable for an upscale Italian restaurant, by implementing a machine learning model that uses Clustering; specifically the k-means algorithm.

Clustering is the task of identifying similar instances and assigning them to *clusters*, or groups of similar instances. In this method the task was to cluster neighborhoods that have established Italian restaurants. As mentioned earlier. the Company has been successful in the past by launching one of its restaurants in locations where there are already similar type businesses established.

I used a Jupyter Notebook environment and the Python programming language to write the required code and textual references needed to arrive at my final conclusions.
The complete Jupyter Notebook is available for review at:
https://github.com/nnoplock/ecdLabs/blob/master/The_Battle_of_the_Neighborhoods_Notebook_Noplock.ipynb

Several Python libraries were used including

- *Numpy*, a library to handle data in a vectorized manners
- *Pandas*, a library for data analysis
- *Json*, a library to handle JSON files
- *Geopy*, (geopy.geocoders <Nominatim>) to convert an address into Latitude and Longitude values
- *Json_normalize*, to transform JSON file data into a pandas dataframe
- *Matplotlib*, for plotting results
- *k-means*, for clustering
- *Folium*, a map rendering library

Research Methods & Data Collection

The primary dataset, that provides New York City borough and neighborhood information, is available from the New York University Spatial Data Repository at https://cocl.us/new_york_dataset

I also obtained location data from FourSquare and demographics from the US Census bureau. The New York Dataset provides Borough, Neighborhood, and location coordinates for all of New York City. I initially intended to single out one or perhaps two suitable New York Boroughs (there are 5 to select from) and then merge FourSquare venue data in order to determine New York neighborhoods (there are 306) that have established Italian restaurants. I have been instructed to first review Manhattan New York and have done so for this report.

Data and Model

The data model was created through a series of steps that included acquiring and preprocessing the needed data and then converting and merging the data in order to produce the final analysis.

1.) I created the object **`newyork_data`** and stored the *new_york_dataset* json data in that object. Below is a sample of the json data that contains New York borough and neighborhood features. This sample shows data for the Bronx borough and the neighborhood, Wakefield.


```
{'type': 'FeatureCollection',
 'totalFeatures': 306,
 'features': [{'type': 'Feature',
   'id': 'nyu_2451_34572.1',
   'geometry': {'type': 'Point',
    'coordinates': [-73.84720052054902, 40.89470517661]},
   'geometry_name': 'geom',
   'properties': {'name': 'Wakefield',
    'stacked': 1,
    'annoline1': 'Wakefield',
    'annoline2': None,
    'annoline3': None,
    'annoangle': 0.0,
    'borough': 'Bronx',
    'bbox': [-73.84720052054902,
     40.89470517661,
     -73.84720052054902,
     40.89470517661]}},
```

2.) A json data structure is essentially nested Python dictionaries that can easily be transformed into a pandas dataframe. In this step I completed several actions to fill a dataframe, named **`neighborhoods`**, with the json data.

At this point, and as instructed by my boss, I was only interested in continuing with the data from the borough of Manhattan and its neighborhoods. As such, I've sliced the original dataframe, 'neighborhoods' and created a new dataframe **`manhattan_data`**.

This is a sample of the **`manhattan_data`** dataframe's contents showing the first five neighborhoods in Manhattan

	Borough	Neighborhood	Latitude	Longitude
0	Manhattan	Marble Hill	40.876551	-73.910660
1	Manhattan	Chinatown	40.715618	-73.994279
2	Manhattan	Washington Heights	40.851903	-73.936900
3	Manhattan	Inwood	40.867684	-73.921210
4	Manhattan	Hamilton Heights	40.823604	-73.949688

3.) As mentioned earlier I merged Foursquare location data with New York City data. Requesting data from Foursquare requires setting up an account that includes Foursquare API credentials. I created the GET request URL in order to request venue data, (within a 500 meter radius from each neighborhood's long/lat coordinates) from Foursquare for all neighborhoods in Manhattan.

4.) The following additional actions were completed in order to establish a structure suitable for the clustering of the restaurant data.

- Get categorical data from Foursquare.
- Determine the number of each of the venue types (category) located in each neighborhood.
- Get the statistical mean of the frequency of occurrence of each category in order to group rows by neighborhood.
- Calculate the top 5 most common venues for each neighborhood.

- Create a function, that returns the most common venues and sorts the venues in descending order.

This is a sample of the **`neighborhoods_venues_sorted`** `dataframe`

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue
0	Battery Park City	Park	Hotel	Coffee Shop	Boat or Ferry	Clothing Store	Gym	Memorial Site	Pizza Place	BBQ Joint
1	Carnegie Hill	Coffee Shop	Café	French Restaurant	Wine Shop	Bookstore	Gym / Fitness Center	Bakery	Gym	Japanese Restaurant
2	Central Harlem	African Restaurant	Seafood Restaurant	Bar	American Restaurant	French Restaurant	Chinese Restaurant	Spa	Bookstore	Boutique
3	Chelsea	Coffee Shop	Bakery	American Restaurant	Italian Restaurant	Hotel	Ice Cream Shop	French Restaurant	Art Gallery	Cupcake Shop
4	Chinatown	Chinese Restaurant	Cocktail Bar	American Restaurant	Hotpot Restaurant	Optical Shop	Salon / Barbershop	Bakery	Dessert Shop	Vietnamese Restaurant

5.) The next several actions uses k-means modeling to cluster (5 clusters) discriminating venue categories that distinguish each cluster. The clusters will be analyzed for suitable locations for a new Italian Restaurant.

- Created a new dataframe named, **`manhattan_merged`**, that includes the cluster as well as the top 10 venues for each neighborhood.
- Map and visualize the resulting clusters.

The following result sets are dataframe clusters that show content for Manhattan neighborhoods and the neighborhood's top five most common venues. While inspecting each cluster keep in mind that the Company is seeking to establish an upscale Italian restaurant in neighborhoods where there are already established Italian restaurants and eateries.

Cluster 1

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
9	Yorkville	Italian Restaurant ★	Gym	Coffee Shop	Bar	Pizza Place
10	Lenox Hill	Italian Restaurant ★	Coffee Shop	Sushi Restaurant	Pizza Place	Cocktail Bar
13	Lincoln Square	Café	Theater	Italian Restaurant	Plaza	Performing Arts Venue
14	Clinton	Theater	Italian Restaurant	Gym / Fitness Center	Spa	American Restaurant
15	Midtown	Hotel	Coffee Shop	Sporting Goods Shop	Clothing Store	Steakhouse
16	Murray Hill	Sandwich Place	Coffee Shop	American Restaurant	Hotel	Japanese Restaurant
25	Manhattan Valley	Indian Restaurant	Bar	Mexican Restaurant	Coffee Shop	Chinese Restaurant
28	Battery Park City	Park	Hotel	Coffee Shop	Boat or Ferry	Clothing Store
29	Financial District	Coffee Shop	Bar	American Restaurant	Cocktail Bar	Pizza Place
30	Carnegie Hill	Coffee Shop	Café	French Restaurant	Wine Shop	Bookstore
32	Civic Center	Coffee Shop	Hotel	Gym / Fitness Center	Cocktail Bar	Spa
33	Midtown South	Korean Restaurant	Japanese Restaurant	Hotel	Dessert Shop	Coffee Shop
34	Sutton Place	Italian Restaurant ★	Gym / Fitness Center	Gym	Coffee Shop	American Restaurant
38	Flatiron	Yoga Studio	Spa	Japanese Restaurant	American Restaurant	Gym / Fitness Center
39	Hudson Yards	American Restaurant	Gym / Fitness Center	Italian Restaurant	Hotel	Café

Cluster 2

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
37	Stuyvesant Town	Boat or Ferry	Bar	Park	Heliport	Gym / Fitness Center

Continued on next page

Cluster 3

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
1	Chinatown	Chinese Restaurant	Cocktail Bar	American Restaurant	Hotpot Restaurant	Optical Shop
5	Manhattanville	Coffee Shop	Seafood Restaurant	Italian Restaurant	Deli / Bodega	Mexican Restaurant
6	Central Harlem	African Restaurant	Seafood Restaurant	Bar	American Restaurant	French Restaurant
8	Upper East Side	Coffee Shop	Italian Restaurant	Exhibit	Bakery	Gym / Fitness Center
12	Upper West Side	Italian Restaurant	Wine Bar	Bar	Bakery	Café
17	Chelsea	Coffee Shop	Bakery	American Restaurant	Italian Restaurant	Hotel
18	Greenwich Village	Italian Restaurant	Sushi Restaurant	Café	Clothing Store	Indian Restaurant
19	East Village	Bar	Pizza Place	Ice Cream Shop	Cocktail Bar	Mexican Restaurant
21	Tribeca	American Restaurant	Italian Restaurant	Park	Spa	Café
22	Little Italy	Bakery	Café	Bubble Tea Shop	Italian Restaurant	Mediterranean Restaurant
23	Soho	Clothing Store	Italian Restaurant	Boutique	Coffee Shop	Shoe Store
24	West Village	Italian Restaurant	New American Restaurant	Jazz Club	American Restaurant	Park
27	Gramercy	Italian Restaurant	Bar	American Restaurant	Bagel Shop	Pizza Place
31	Noho	Italian Restaurant	Pizza Place	Hotel	French Restaurant	Coffee Shop
35	Turtle Bay	Italian Restaurant	Sushi Restaurant	Coffee Shop	Wine Bar	Steakhouse

Cluster 4

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Marble Hill	Sandwich Place	Gym	American Restaurant	Coffee Shop	Yoga Studio

Cluster 5

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
2	Washington Heights	Café	Bakery	Grocery Store	Mobile Phone Shop	Chinese Restaurant
3	Inwood	Café	Mexican Restaurant	Restaurant	Pizza Place	Lounge
4	Hamilton Heights	Pizza Place	Café	Coffee Shop	Deli / Bodega	Mexican Restaurant
7	East Harlem	Mexican Restaurant	Thai Restaurant	Bakery	Pizza Place	Latin American Restaurant
11	Roosevelt Island	Park	Bus Line	Residential Building (Apartment / Condo)	Scenic Lookout	Sandwich Place
20	Lower East Side	Chinese Restaurant	Coffee Shop	Café	Bakery	Pizza Place
26	Morningside Heights	Park	Coffee Shop	American Restaurant	Bookstore	Burger Joint
36	Tudor City	Park	Café	Mexican Restaurant	Deli / Bodega	Diner

Results

After inspecting the 5 clusters it has been determined, at a very high level, that only two clusters include neighborhoods where the Italian Restaurant venue is the "1st Most Common Venue" as marked with a blue star ★.

Cluster 1 neighborhoods	Cluster 2 neighborhoods
Lincoln Square	Upper East Side
Greenwich Village	Lenox Hill
West Village	Upper West Side
Gramercy	
Noho	

Additional analysis and due diligence will be applied to each of the neighborhoods listed above in order to determine suitable locations for a new Italian Restaurant in Manhattan New York.

The table below shows that I added US Census data that provides population, income and housing demographics as additional decision making criteria.

Cluster 1 neighborhoods	Population	Median Household Income	People Above Poverty Level	Median Age	Total Households
Lincoln Square	49,957	120,538.00	44,153.00	40	26,795
Greenwich Village	30,591	119,366.00	25,101.00	36	16,196
West Village	30,738	126,601	28,194	39.4	18,229
Gramercy	87,783	115,027	77,345	36.3	9,197
Noho	5,579	110,584	4,389	33	2,769
Cluster 2 neighborhoods					
Upper East Side	199,437	126,258	189,465	40	104973
Lenox Hill	88,459	132,260	81,571	40	46,978
Upper West Side	191,222	115,341	175,924	42	95,868

Data Sources

New York data set

https://cocl.us/new_york_dataset

Foursquare location data

https://foursquare.com/

United States Census data

https://data.census.gov/cedsci/
https://census.missouri.edu/
https://censusreporter.org/search/
https://datausa.io/