Analyzing a suitable location for a new restaurant near Manhattan

Introduction/Business Problem

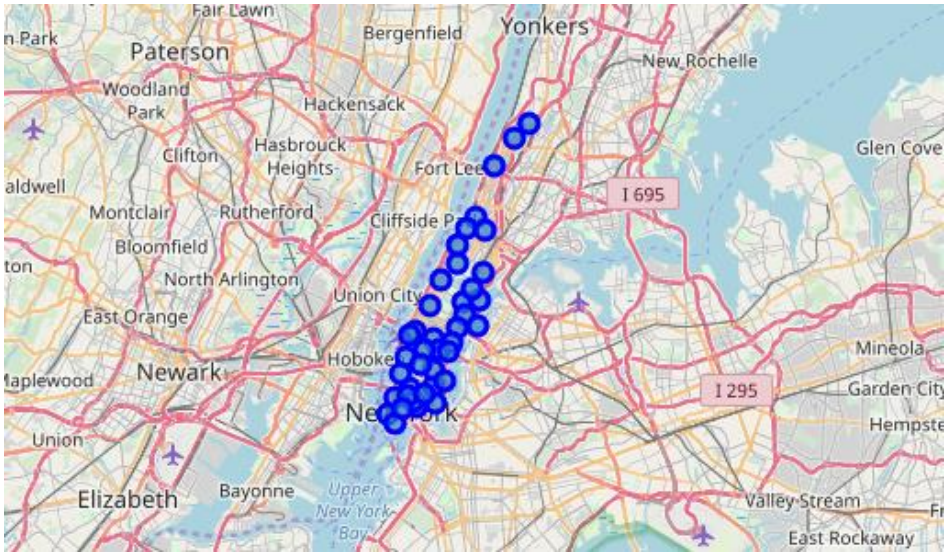
According to the *World Population Review* 2019: Manhattan is the most densely populated of the five boroughs in New York. It has an estimated population of 1.63 million people, all living in an area of just 23 square miles. This gives Manhattan a population density of 70,826 people per square mile, or 27,346 per square mile, which is denser than any city in the country. It is also one of the most densely populated areas on earth. Manhattan has the highest cost of living in the United States and the only county in the country with a per capita income of more than $100,000. Being heart to several iconic locations such as the Time Square, NASDAQ, the New York Stock Exchange and Wall Street, it is one of the commercial, financial and cultural centers in the world. Given all the above information, a client decides to open a restaurant near Manhattan. This project aims to find a suitable location for the new restaurant.

Data Acquisition and Preparation

This analysis mainly uses a data set freely available on *https://geo.nyu.edu/catalog/nyu_2451_34572*, which captures the neighborhood names and latitude-longitude information for all 5 boroughs and 306 neighborhoods of New York. An extract from the relevant data for the Manhattan Borough, extracted to a pandas dataframe is as follows:

	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
5	Manhattan	Manhattanville	40.816934	-73.957385
6	Manhattan	Central Harlem	40.815976	-73.943211

A visual representation of the neighbourhoods considered are shown below, using a Folium map:

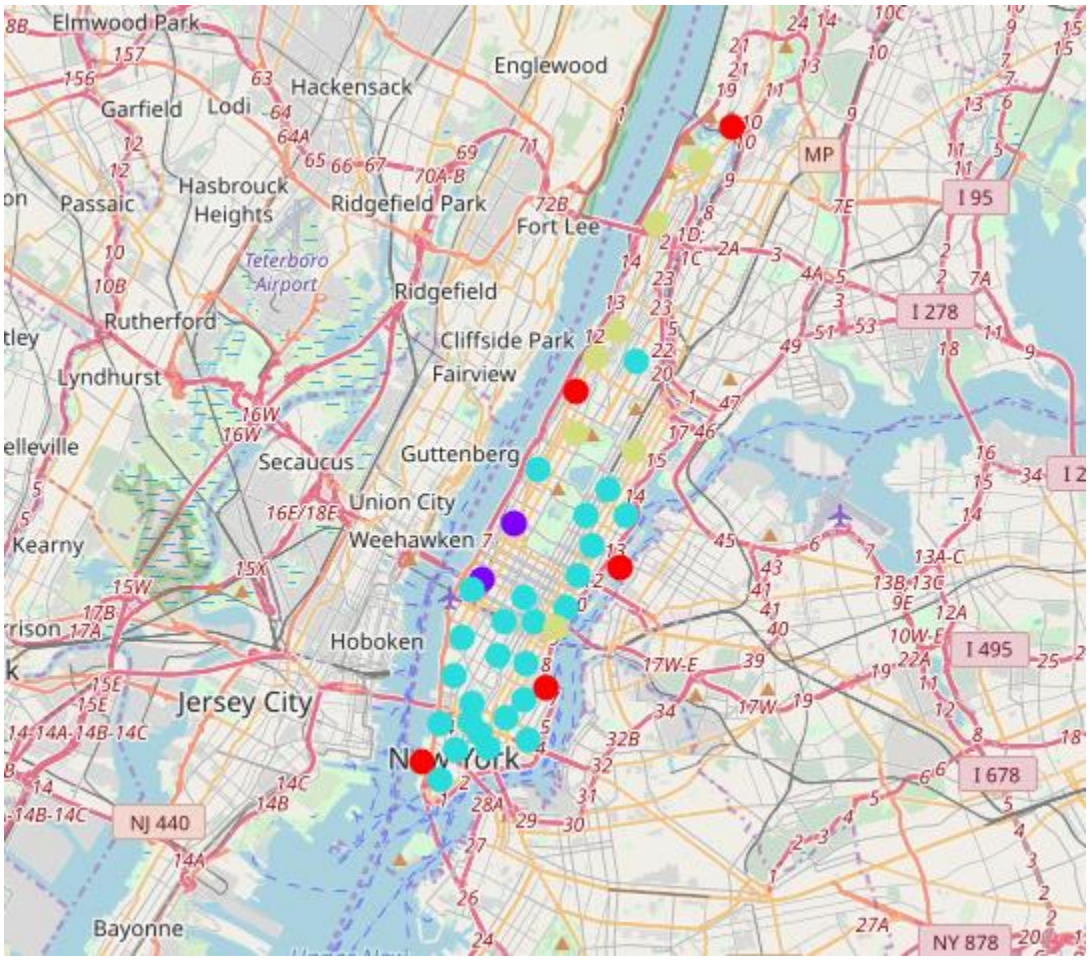


Following the initial data acquisition step above, *Foursquare* API was used to extract information on the popular venues in the neighborhoods of Manhattan which were processed into a pandas dataframe displaying the top (most frequent) 5 venues for each neighbourhood. An extract of the resulting dataframe is shown below:

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Battery Park City	Park	Coffee Shop	Memorial Site	BBQ Joint	Food Court
1	Carnegie Hill	Pizza Place	Bookstore	Coffee Shop	Spa	Italian Restaurant
2	Central Harlem	African Restaurant	Gym / Fitness Center	American Restaurant	Bar	Public Art
3	Chelsea	Ice Cream Shop	Hotel	Nightclub	Cupcake Shop	Seafood Restaurant
4	Chinatown	Chinese Restaurant	Salon / Barbershop	Spa	Sandwich Place	Cocktail Bar

Data Analysis

Following the data preparation phase, K-means clustering (a popular unsupervised machine learning algorithms) was used to divide the neighbourhoods of Manhattan into 4 clusters based on the similarity of venue occurrences. The visualization map and the tabulations of the 4 clusters are as follows:



Cluster 1 (red circles)

	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	Coffee Shop	Yoga Studio	Clothing Store	Seafood Restaurant
11	Roosevelt Island	Park	Deli / Bodega	Coffee Shop	Sandwich Place	Noodle House
26	Morningside Heights	Bookstore	American Restaurant	Coffee Shop	Park	Tennis Court
28	Battery Park City	Park	Coffee Shop	Memorial Site	BBQ Joint	Food Court
37	Stuyvesant Town	Park	Boat or Ferry	Bar	Playground	Cocktail Bar

Cluster 2 (purple circles)

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
13	Lincoln Square	Theater	Concert Hall	Plaza	Indie Movie Theater	Opera House
14	Clinton	Theater	Gym / Fitness Center	Wine Shop	Indie Theater	Lounge

Cluster 3 (light blue circles)

	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	Salon / Barbershop	Spa	Sandwich Place	Cocktail Bar
6	Central Harlem	African Restaurant	Gym / Fitness Center	American Restaurant	Bar	Public Art
8	Upper East Side	Italian Restaurant	Hotel	Boutique	Bakery	American Restaurant
9	Yorkville	Italian Restaurant	Gym	Deli / Bodega	Coffee Shop	Café
10	Lenox Hill	Burger Joint	Gym	Gym / Fitness Center	Bakery	Turkish Restaurant
12	Upper West Side	Seafood Restaurant	Bakery	Cosmetics Shop	Italian Restaurant	Bar
15	Midtown	Clothing Store	Sporting Goods Shop	Bookstore	Cocktail Bar	Coffee Shop
16	Murray Hill	Hotel	Japanese Restaurant	Coffee Shop	Burger Joint	Italian Restaurant
17	Chelsea	Ice Cream Shop	Hotel	Nightclub	Cupcake Shop	Seafood Restaurant
18	Greenwich Village	Italian Restaurant	Sushi Restaurant	Café	Cosmetics Shop	Clothing Store
19	East Village	Pizza Place	Bar	Coffee Shop	Vietnamese Restaurant	Dessert Shop

Cluster 4 (light yellow circles)

	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	Coffee Shop	Mobile Phone Shop	Wine Shop	Bakery
3	Inwood	Café	Mexican Restaurant	Pizza Place	American Restaurant	Park
4	Hamilton Heights	Café	Mexican Restaurant	Coffee Shop	Yoga Studio	Sushi Restaurant
5	Manhattanville	Coffee Shop	Italian Restaurant	Mexican Restaurant	Seafood Restaurant	Sushi Restaurant
7	East Harlem	Mexican Restaurant	Bakery	Deli / Bodega	Latin American Restaurant	Thai Restaurant
25	Manhattan Valley	Indian Restaurant	Coffee Shop	Pizza Place	Playground	Bar
36	Tudor City	Park	Deli / Bodega	Mexican Restaurant	Café	Thai Restaurant

Conclusion

From the above tabulations, it is evident that the two neighborhoods Lincoln Square and Clinton (in cluster 2) provide the ideal locations for a new restaurant, mainly due to the ample availability in leisure activity locations such as theatres (after visiting which, people are highly likely to dine outside), and due to the lowest competition from other restaurants.