The battle of Neighborhoods

Problem background :

New York City (NYC), often called simply New York, is the most populous city in the United States. With an estimated 2019 population of 8,336,817 distributed over about 302.6 square miles, New York City is also the most densely populated major city in the United States. Located at the southern tip of the U.S. state of New York, the city is the center of the New York metropolitan area, the largest metropolitan area in the world by urban landmass. With almost 20 million people in its metropolitan statistical area and approximately 23 million in its combined statistical area, it is one of the world's most populous megacities.

Situated on one of the world's largest natural harbors, New York City is composed of five boroughs, each of which is a county of the State of New York. The five boroughs—Brooklyn, Queens, Manhattan, the Bronx, and Staten. The city and its metropolitan area constitute the premier gateway for legal immigration to the United States. As many as 800 languages are spoken in New York, making it the most linguistically diverse city in the world. New York is home to more than 3.2 million residents born outside the United States, the largest foreign-born population of any city in the world as of 2016. As of 2019, the New York metropolitan area is estimated to produce a gross metropolitan product (GMP) of $2.0 trillion. (Reference: Wikipedia)

Problem description:

The objective of this capstone project is not only to analyze/select the best locations in the city of new york for tourists, but also to give suggestions to those who want to open a new restaurant in new york city, based on the data of different neighborhoods in the city.

Based on the above info, we will answer three questions below:
  1. What the most visited venus in New York City?
  2. Which neighborhood has the most venues?
  3. What the best investment choices in new york city if someone wants to open a restaurant?

By answering the above three questions, it will help tourists or investors to know what place is better to visit, invest, or to start their own business based on the venues' information. This will also help people to make decisions if they are thinking about migrating to New York City. In short, the target for this project is basically everyone who is exploring different places or similar places.

Description of the data

The main data that will use in the project is from below link:
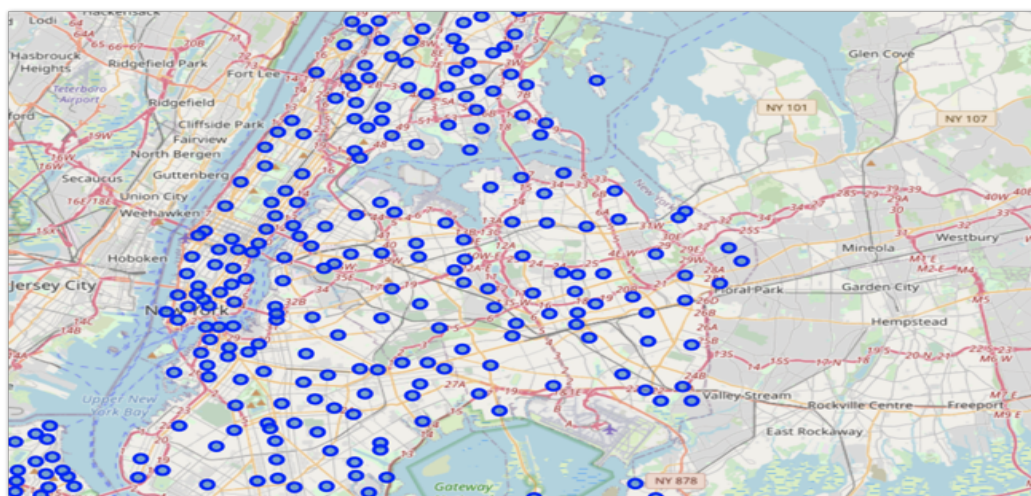https://maps.princeton.edu/catalog/nyu-2451-34572.
we will use the below data:
1. List of neighborhoods in new york city.
2. Latitude and longitude coordinates of those neighborhoods
3. Specific venue data and mainly related to restaurants

Data We Need

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Wakefield	40.894705	-73.847201	Lollipops Gelato	40.894123	-73.845892	Dessert Shop
1	Wakefield	40.894705	-73.847201	Rite Aid	40.896649	-73.844846	Pharmacy
2	Wakefield	40.894705	-73.847201	Walgreens	40.896528	-73.844700	Pharmacy
3	Wakefield	40.894705	-73.847201	Carvel Ice Cream	40.890487	-73.848568	Ice Cream Shop
4	Wakefield	40.894705	-73.847201	Dunkin'	40.890459	-73.849089	Donut Shop
5	Wakefield	40.894705	-73.847201	Subway	40.890468	-73.849152	Sandwich Place
6	Wakefield	40.894705	-73.847201	Central Deli	40.896728	-73.844387	Deli / Bodega
7	Wakefield	40.894705	-73.847201	Koss Quick Wash	40.891281	-73.849904	Laundromat
8	Co-op City	40.874294	-73.829939	Capri II Pizza	40.876374	-73.829940	Pizza Place
9	Co-op City	40.874294	-73.829939	Rite Aid	40.870345	-73.828302	Pharmacy

Use Folium to visualize the data

Implement Foursquare API to explore

Use Foursquare API to get the nearest venue info, so that we can use them to form a cluster. And, the results will show the corresponding coordinates with geoinformation as well.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Wakefield	40.894705	-73.847201	Lollipops Gelato	40.894123	-73.845892	Dessert Shop
1	Wakefield	40.894705	-73.847201	Rite Aid	40.896649	-73.844846	Pharmacy
2	Wakefield	40.894705	-73.847201	Walgreens	40.896528	-73.844700	Pharmacy
3	Wakefield	40.894705	-73.847201	Carvel Ice Cream	40.890487	-73.848568	Ice Cream Shop
4	Wakefield	40.894705	-73.847201	Dunkin'	40.890459	-73.849089	Donut Shop
5	Wakefield	40.894705	-73.847201	Subway	40.890468	-73.849152	Sandwich Place
6	Wakefield	40.894705	-73.847201	Central Deli	40.896728	-73.844387	Deli / Bodega
7	Wakefield	40.894705	-73.847201	Koss Quick Wash	40.891281	-73.849904	Laundromat
8	Co-op City	40.874294	-73.829939	Capri II Pizza	40.876374	-73.829940	Pizza Place
9	Co-op City	40.874294	-73.829939	Rite Aid	40.870345	-73.828302	Pharmacy

Also, we collect the most common venues from each neighborhood

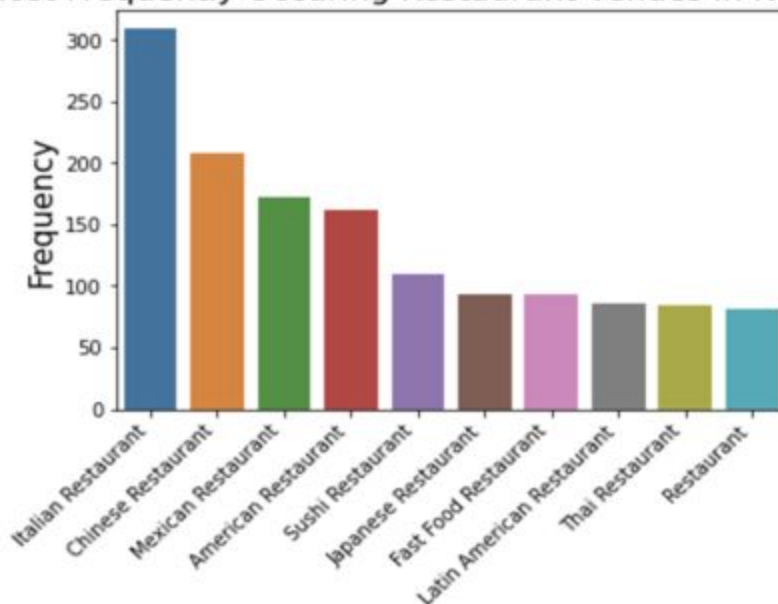|	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Allerton	Pizza Place	Supermarket	Chinese Restaurant	Deli / Bodega	Spanish Restaurant
1	Annadale	Pizza Place	Sushi Restaurant	Bakery	Liquor Store	Pharmacy
2	Arden Heights	Deli / Bodega	Pharmacy	Coffee Shop	Business Service	Pizza Place
3	Arlington	Bus Stop	Intersection	Deli / Bodega	American Restaurant	Women's Store
4	Arrochar	Bus Stop	Italian Restaurant	Pizza Place	Deli / Bodega	Liquor Store
5	Arverne	Surf Spot	Metro Station	Sandwich Place	Wine Shop	Thai Restaurant
6	Astoria	Bar	Hookah Bar	Middle Eastern Restaurant	Café	Bakery
7	Astoria Heights	Italian Restaurant	Playground	Bakery	Business Service	Hostel
8	Auburndale	Italian Restaurant	Athletics & Sports	Noodle House	Mobile Phone Shop	Supermarket
9	Bath Beach	Chinese Restaurant	Pizza Place	Cantonese Restaurant	Pharmacy	Gas Station
10	Battery Park City	Coffee Shop	Park	Hotel	Clothing Store	Memorial Site
11	Bay Ridge	Italian Restaurant	Spa	Pizza Place	Bagel Shop	American Restaurant
12	Bay Terrace	Clothing Store	Supermarket	Donut Shop	Cosmetics Shop	Shoe Store
13	Baychester	Donut Shop	Fried Chicken Joint	Discount Store	Bus Station	Supermarket
14	Bayside	Bar	Pizza Place	Indian Restaurant	Sushi Restaurant	Italian Restaurant
15	Bayswater	Construction & Landscaping	Tennis Court	Playground	Women's Store	Fast Food Restaurant
16	Bedford Park	Mexican Restaurant	Chinese Restaurant	Diner	Deli / Bodega	Pizza Place
17	Bedford Stuyvesant	Coffee Shop	Café	Deli / Bodega	Bar	Pizza Place
18	Beechhurst	Italian Restaurant	Donut Shop	Supermarket	Boutique	Optical Shop
19	Bellaire	Deli / Bodega	Italian Restaurant	Breakfast Spot	Gym	Grocery Store

Methodology

On the other hand, we create a data frame that only focuses on restaurants. And, by applying seaborn and matplotlib libraries to visualize the top 10 most frequently occurring restaurants venues in New York City.

```
print (ny_Venues_only_restaurant['Venue Category'].value_counts())
```

```
Italian Restaurant               309
Chinese Restaurant               208
Mexican Restaurant               172
American Restaurant              162
Sushi Restaurant                 110
Japanese Restaurant               93
Fast Food Restaurant              93
Latin American Restaurant         86
Thai Restaurant                   84
Restaurant                        81
Caribbean Restaurant              77
Korean Restaurant                 76
Seafood Restaurant                74
Spanish Restaurant                72
French Restaurant                 71
Indian Restaurant                 65
Asian Restaurant                  46
Mediterranean Restaurant          46
New American Restaurant           40
Vietnamese Restaurant             39
Vegetarian / Vegan Restaurant     38
Greek Restaurant                  36
Middle Eastern Restaurant         34
Ramen Restaurant                  27
```


10 Most Frequently Occuring Restaurant Venues in New York City

Eventually, group the neighborhoods into different clusters by applying the K-means algorithm to get the final results.

	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Bronx	Wakefield	40.894705	-73.847201	4	Pharmacy	Deli / Bodega	Donut Shop	Sandwich Place	Dessert Shop
1	Bronx	Co-op City	40.874294	-73.829939	4	Fast Food Restaurant	Bus Station	Pizza Place	Post Office	Basketball Court
2	Bronx	Eastchester	40.887556	-73.827806	4	Caribbean Restaurant	Bus Station	Diner	Deli / Bodega	Platform
3	Bronx	Fieldston	40.895437	-73.905643	0	Medical Supply Store	Bus Station	River	Plaza	Women's Store
4	Bronx	Riverdale	40.890834	-73.912585	0	Park	Bank	Bus Station	Gym	Playground

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
3	Fieldston	Medical Supply Store	Bus Station	River	Plaza	Women's Store
4	Riverdale	Park	Bank	Bus Station	Gym	Playground
9	Williamsbridge	Bar	Nightclub	Caribbean Restaurant	Soup Place	Fast Food Restaurant
12	City Island	Seafood Restaurant	Thrift / Vintage Store	Italian Restaurant	Arts & Crafts Store	Spanish Restaurant
18	West Farms	Bus Station	Park	Donut Shop	Playground	Cuban Restaurant
22	Port Morris	Furniture / Home Store	Bar	Distillery	Spanish Restaurant	Brewery
28	Throgs Neck	Italian Restaurant	Pizza Place	Asian Restaurant	Baseball Field	Deli / Bodega
29	Country Club	Sandwich Place	Italian Restaurant	Flea Market	Playground	Athletics & Sports
39	Edgewater Park	Italian Restaurant	Deli / Bodega	Pizza Place	Coffee Shop	Asian Restaurant
41	Olinville	Caribbean Restaurant	Supermarket	Fried Chicken Joint	Convenience Store	Deli / Bodega

Discussion/Conclusion

Overall New York City is a diverse city. Based on the data above,
Italian Restaurant, Coffee shops, Pizza places are the most common restaurants in the city.

On the other hand, there are a few interesting things:
1. In neighborhood 'Wingate', 'Holliswood', and 'New Lots', it seems like people really love fried chicken
2. Boba tea shop is the most common venue in neighborhood 'Whitestone'
3. Indian restaurants are the most common restaurant in neighborhood 'Floral Park'

In conclusion, this project involved lots of data cleaning, data wrangling, machine learning, and data visualization. By using these data, it can help potential business owners to make decisions on which locations they can invest in.