

The Battle of Neighborhoods Report

Introduction

Background

New York City is one of the most popular cities in the United States, and it comprises 5 boroughs. This international city is the leading cultural center and economic center of the world. It hosts many landmarks, including Time Square, Central Park, Empire State Building, and etc. Manhattan, the core of New York City, is the most densely populated borough in New York City, and it serves as the city's economic and administrative center.

Problem

Since Manhattan is the core borough of the New York City, many cultures will be combined in this city, and it appears food diversity as well as people diversity. The market is also highly competitive, so this analysis will provide a good understanding of food market environment which will show which type of food is most popular in Manhattan, and it will also help the new business starters related to restaurant/food identify which type of restaurant/food they will open. This will help in reduction of risks from competitive power.

Interest

Obviously, starters who want to open a restaurant in Manhattan will be interested in understanding the food market environment for competitive power. And others who travel to Manhattan and who want to get touch in local culture can also be very interested in the most popular types of food in Manhattan.

Data

I will use the following dataset to show the analysis of food market environment in Manhattan: First, I will use the New York City Dataset, with the link below:

https://geo.nyu.edu/catalog/nyu_2451_34572

This dataset shows the neighborhoods that exist in each of five boroughs in New York with its latitude and longitude, including Manhattan.

The view of the data would be:

```
{'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]}},
```

Second, I will use the Foursquare API database, with the link below:

<https://developer.foursquare.com/doc>

I searched the food category in Foursquare API dataset. Venues retrieved from all the neighborhoods in New York City are categorized into ‘American Restaurant’, ‘Butcher’, ‘Chinese Restaurant’, ‘Deli/Bodega’, ‘Food Court’, ‘Food Truck’, Frozen Yogurt Shop’, ‘Gourmet Shop’, ‘Grocery Store’, ‘Italian Restaurant’, ‘Museum’, ‘Organic Grocery’, ‘Other Repair Shop’, ‘Restaurant’, ‘Southern/Soul Food Restaurant’, and ‘Supermarket.’

I will use groupby function to only focus on Manhattan data, so the view of the data for Manhattan (the first 5 lines) would be:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Marble Hill	40.876551	-73.91066	PeraBell Food Bar	40.765781	-73.013731	American Restaurant
1	Marble Hill	40.876551	-73.91066	Aldi Food Market	40.778311	-73.033230	Supermarket
2	Marble Hill	40.876551	-73.91066	Delfiore Pizza & Food Co.	40.765692	-73.013344	Italian Restaurant
3	Marble Hill	40.876551	-73.91066	Best Meal Chinese Food	40.765382	-73.013084	Chinese Restaurant
4	Marble Hill	40.876551	-73.91066	Smith Haven Mall Food Court	40.863369	-73.129668	Food Court

Methodology

The New York Dataset

In order to see the neighborhoods of each of five boroughs in New York City, I download and explore the dataset with a total of 5 boroughs and 306 neighborhoods, with the latitude and longitude coordinates of each neighborhoods.

This dataset is download using the URL like provided above. And then the dataset .json file is downloaded, and I explore the data to understand the structure of the data and get the overview of the data.

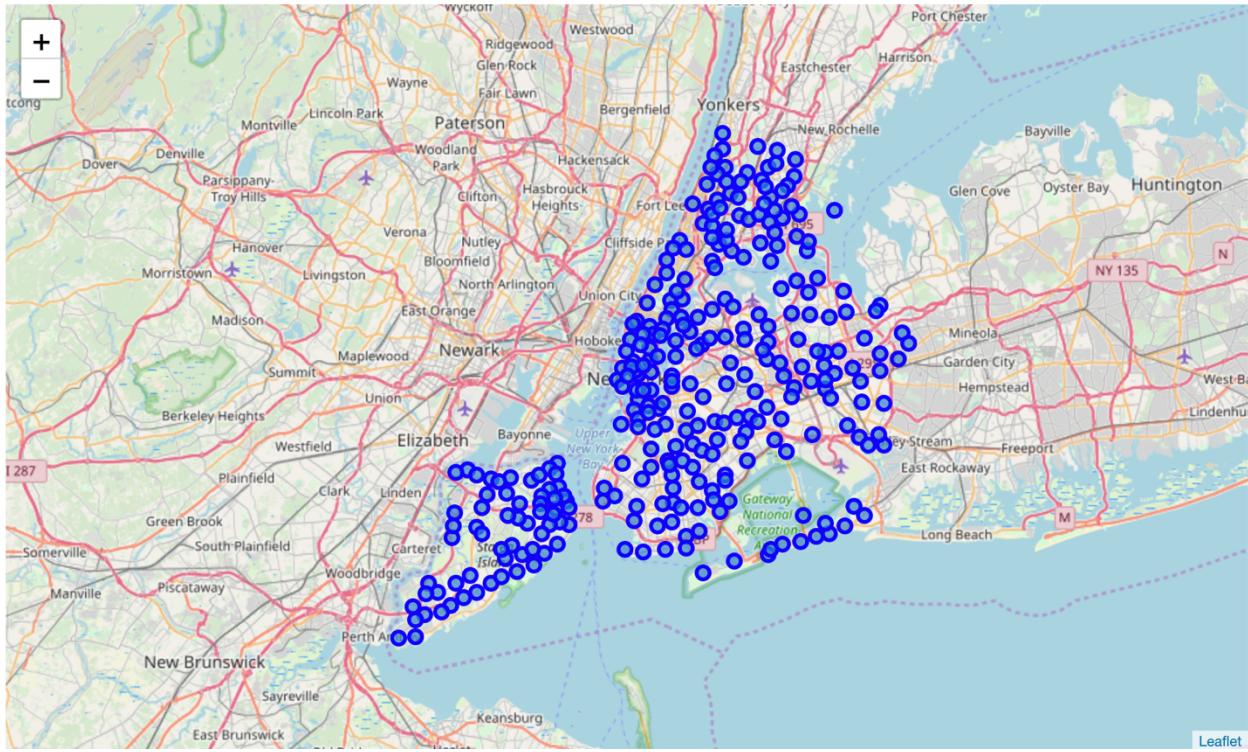
Then, the data is transformed into a pandas dataframe, and I loop through the data to fill the dataframe one row at a time. After doing these steps, the dataframe is created with the columns ‘Borough,’ ‘Neighborhood,’ ‘Latitude,’ and ‘Longitude.’

The database is shown as below:

	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

Then, I used geopy library to get the latitude and longitude coordinate of New York City, which is (40.7127281, -74.0060152).

Then, the folium help to visualize the data by creating the map to show the neighborhoods in New York, and the map is shown as below:

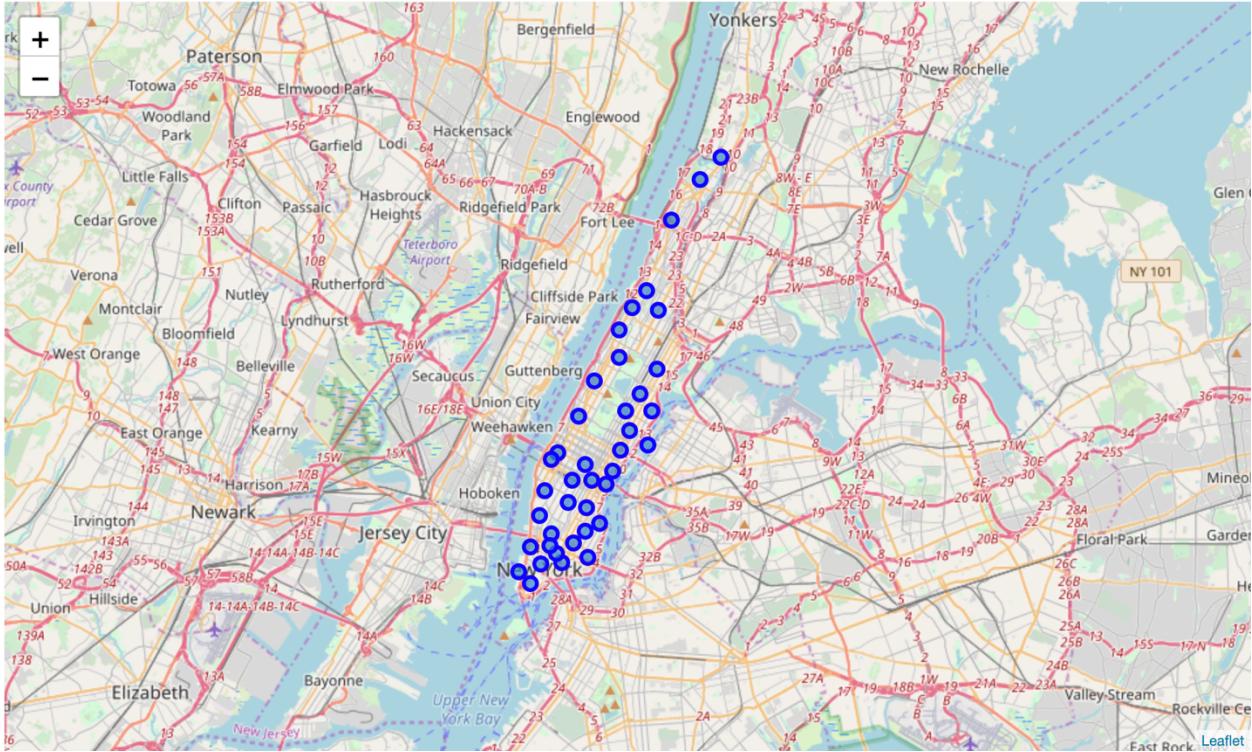


Focus on Manhattan

Since my focus is on Manhattan, I cleaned the data and only show the data related to the Manhattan, which means ‘Borough’ equals to ‘Manhattan.’ The dataframe now is shown as:

Borough	Neighborhood	Latitude	Longitude
0	Manhattan	Marble Hill	40.876551
1	Manhattan	Chinatown	40.715618
2	Manhattan	Washington Heights	40.851903
3	Manhattan	Inwood	40.867684
4	Manhattan	Hamilton Heights	40.823604

Then I used the same step as above to create the map of Manhattan:



Foursquare API

First, I created the client_id and client_secret in Foursquare website, and then I defined the ‘client_id,’ ‘client_secret,’ ‘version,’ and ‘limit’ to get access of API.

Then, I created the API request URL, made the GET request, and only returned the information I needed for each venue. And then, my dataframe turns out into a new dataframe with only Manhattan related information, with the columns of ‘Neighborhood,’ ‘Neighborhood Latitude,’ ‘Neighborhood Longitude,’ ‘Venue,’ ‘Venue Latitude,’ ‘Venue Longitude,’ and ‘Venue Category.’ The dataframe is shown as below:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Marble Hill	40.876551	-73.91066	PeraBell Food Bar	40.765781	-73.013731	American Restaurant
1	Marble Hill	40.876551	-73.91066	Aldi Food Market	40.778311	-73.033230	Supermarket
2	Marble Hill	40.876551	-73.91066	Delfiore Pizza & Food Co.	40.765692	-73.013344	Italian Restaurant
3	Marble Hill	40.876551	-73.91066	Best Meal Chinese Food	40.765382	-73.013084	Chinese Restaurant
4	Marble Hill	40.876551	-73.91066	Smith Haven Mall Food Court	40.863369	-73.129668	Food Court

Then, I counted the number of each type of restaurant for each neighborhood (by using groupby function), and then calculate the occurrence for each type of venue. The table is shown as below:

Neighborhood	American Restaurant	Butcher	Chinese Restaurant	Deli / Bodega	Food Court	Food Truck	Frozen Yogurt Shop	Gourmet Shop	Grocery Store	Italian Restaurant	Museum	Organic Grocery	Other Repair Shop	Restaurant	Southern / Soul Food Restaurant	Supermarket
0	0.066667	0.033333	0.066667	0.033333	0.166667	0.266667	0.033333	0.033333	0.033333	0.033333	0.033333	0.033333	0.033333	0.033333	0.033333	0.066667
1	0.066667	0.033333	0.066667	0.033333	0.166667	0.266667	0.033333	0.033333	0.033333	0.033333	0.033333	0.033333	0.033333	0.033333	0.033333	0.066667
2	0.066667	0.033333	0.066667	0.033333	0.166667	0.266667	0.033333	0.033333	0.033333	0.033333	0.033333	0.033333	0.033333	0.033333	0.033333	0.066667
3	0.066667	0.033333	0.066667	0.033333	0.166667	0.266667	0.033333	0.033333	0.033333	0.033333	0.033333	0.033333	0.033333	0.033333	0.033333	0.066667
4	0.066667	0.033333	0.066667	0.033333	0.166667	0.266667	0.033333	0.033333	0.033333	0.033333	0.033333	0.033333	0.033333	0.033333	0.033333	0.066667

Then, I created the new dataframe with the columns to show the top 10 venues, which tells us the top 10 popular/common type of restaurants opened in Manhattan.

Results

After analyzing the data, we find that the top 10 common venues opened in Manhattan, which these venue types are popular for people to open in Manhattan.

The results of the Cluster 1 analysis are shown below:

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	10th Most Common Venue	
0	Marble Hill	Food Truck	Food Court	Supermarket	Chinese Restaurant	American Restaurant	Southern / Soul Food Restaurant	Restaurant	Other Repair Shop	Organic Grocery	Museum
1	Chinatown	Food Truck	Food Court	Supermarket	Chinese Restaurant	American Restaurant	Southern / Soul Food Restaurant	Restaurant	Other Repair Shop	Organic Grocery	Museum
2	Washington Heights	Food Truck	Food Court	Supermarket	Chinese Restaurant	American Restaurant	Southern / Soul Food Restaurant	Restaurant	Other Repair Shop	Organic Grocery	Museum
3	Inwood	Food Truck	Food Court	Supermarket	Chinese Restaurant	American Restaurant	Southern / Soul Food Restaurant	Restaurant	Other Repair Shop	Organic Grocery	Museum

As we can see, the food truck is the most common venue in Manhattan across different neighborhoods. And food court is the second common venue.

Discussion

After analyzing the data, we find that the top 10 common venues opened in Manhattan, which these venue types are popular for people to open in Manhattan.

These top 10 common venues are food truck, food court, supermarket, Chinese restaurant, American restaurant, southern/soul food restaurant, and etc.

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	10th Most Common Venue	
0	Battery Park City	Food Truck	Food Court	Supermarket	Chinese Restaurant	American Restaurant	Southern / Soul Food Restaurant	Restaurant	Other Repair Shop	Organic Grocery	Museum
1	Carnegie Hill	Food Truck	Food Court	Supermarket	Chinese Restaurant	American Restaurant	Southern / Soul Food Restaurant	Restaurant	Other Repair Shop	Organic Grocery	Museum
2	Central Harlem	Food Truck	Food Court	Supermarket	Chinese Restaurant	American Restaurant	Southern / Soul Food Restaurant	Restaurant	Other Repair Shop	Organic Grocery	Museum
3	Chelsea	Food Truck	Food Court	Supermarket	Chinese Restaurant	American Restaurant	Southern / Soul Food Restaurant	Restaurant	Other Repair Shop	Organic Grocery	Museum
4	Chinatown	Food Truck	Food Court	Supermarket	Chinese Restaurant	American Restaurant	Southern / Soul Food Restaurant	Restaurant	Other Repair Shop	Organic Grocery	Museum

As we can see, these 10 common venues have high competition in the food market in Manhattan, so if people want to open these types of restaurant, their restaurant should have higher taste or highly differentiated/special ideas to attract more customers to come. So if people want to have less competition in the food market, they should open the restaurant that the type is not in these 10 types of common venues.

Conclusion

In this report, I use Clustering to analyze multiple dataset, and provide 10 most common venues (top 10 common types of restaurant) in Manhattan, New York. The results show that food truck and food court are the top 2 most common types.

The results of the project can help the people who want to open the restaurant have an idea about the food market in Manhattan and provide the overview of the competition power for different types of restaurants in Manhattan.

According to the results, I highly suggest people who do not want high competition in the food market avoid opening restaurant related to these 10 types.

References

Notebook for the ‘Applied Data Science Capstone’ course on Coursera