

1- Introduction

1-1 Problem background

New York City (NYC), often called New York (NY), 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 (784 km²), New York is also the most densely populated major city in the United States. Situated on one of the world's largest natural harbors, New York City is composed of five boroughs and 306 neighborhoods, each of which is a county of the State of New York. The five boroughs are Brooklyn, Queens, Manhattan, the Bronx, and Staten Island.

New York City is also considered as a global hub of business and commerce, as a center for banking and finance, retailing, world trade, transportation, tourism, real estate, new media, traditional media, advertising, legal services, accountancy, insurance, theater, fashion, and the arts in the United States. As many as 800 languages are spoken in New York City, 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 a result of city's immigrant history, its food culture includes a variety of international cuisines.

In order to start a new Turkish restaurant in New York City we need to analyze restaurants business environment carefully. The insights derived from our analysis will give good understanding of restaurants business environment in New York City and this will have a major effect on reduce risk of investment and increase Return of Investment (RoI) in this highly competitive market.

1-2 Problem Description

Choosing a location for a new restaurant is one of just a few keys to profitability. There are several factors should be considered when choosing a location for new Turkish restaurant:

- 1- Type of selected area for new restaurant (business areas or residential area)
- 2- Population density: Are there enough people in the selected area to support the business? There need to be enough people who live in or pass through the area regularly to keep the business busy.
- 3- People activity periods in that area (i.e. business area from 9 AM to 5 PM or touristic area can be visited any time of the day)
- 4- Entertainment facilities (cinema, theater, shopping spa, malls, park, garden and playground)
- 5- Transportation
- 6- Competitors in selected location
- 7- Cuisine of the competitors

1-3 Success Criteria

The project aims to select the best location for a new Turkish restaurant in New York City. The selection criteria based on selecting one borough from 5 boroughs in New York City and then explore the neighborhoods of the selected borough to find the most 10 common Venues in each neighborhoods and finally clustering the neighborhoods use the k-means clustering algorithm. After neighborhoods clustering the clusters are explored to select the best cluster as a location for the new Turkish restaurant.

2- Data Acquisition, Cleaning and Exploration

New York City (NYC), often called New York (NY), 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 (784 km²), Situated on one of the world's largest natural harbors, New York City is composed of a total of 5 boroughs and 306 neighborhoods.

in order to analyze New York City to find the best location for Turkish restaurant, the data acquired for this project is a combination of data scraped from a Wikipedia and HTML pages. After accruing and cleaning the data we have four new datasets. We will explore these datasets in order to select the best brought in New York City and then select the best neighborhood in the selected borough to start a new Turkish restaurant.

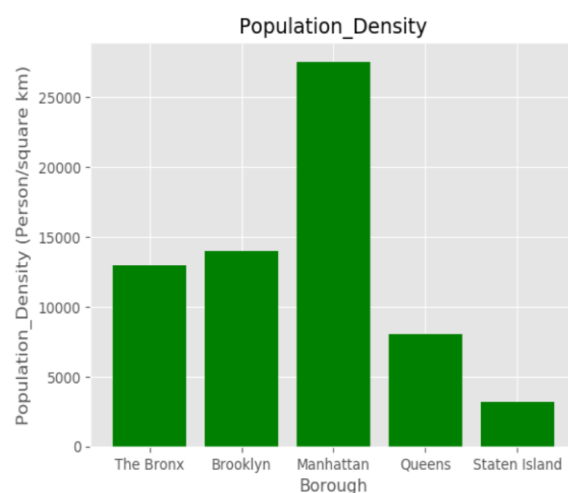
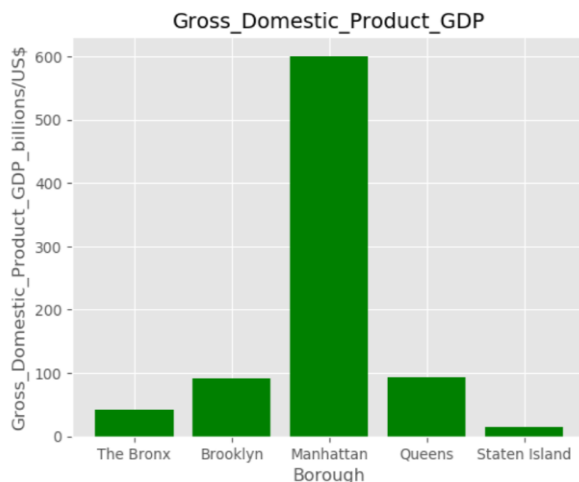
2-1 New York City Data Acquisition, Cleaning and Exploration

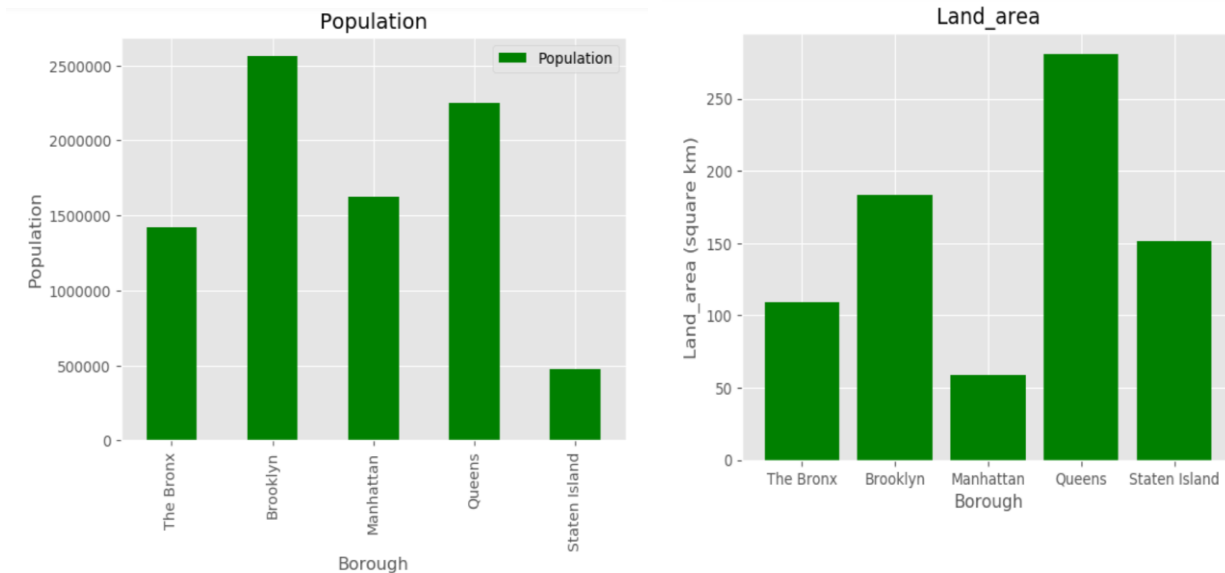
in order to analyze New York City to find the best location for Turkish restaurant, the data acquired for this project is a combination of data scraped from a Wikipedia and HTML pages. After accruing and cleaning the data we have four new datasets. We will explore these datasets in order to select the best brought in New York City and then select the best neighborhood in the selected borough to start a new Turkish restaurant.

First dataset

The first dataset is scraped from a Wikipedia page https://en.wikipedia.org/wiki/New_York_City. The dataset contains the list of New York City boroughs Jurisdiction, population, Gross Domestic Product (GDP), land area and population density.

	Borough	County	Population	Gross Domestic Product billions/US\$	Gross Domestic Product per capita(US\$)	Land area square miles	Land area square km	persons per sq. mile	persons per square km
0	The Bronx	Bronx	1418207	42.695	30100	42.10	109.04	33867	13006
1	Brooklyn	Kings	2559903	91.559	35800	70.82	183.42	36147	13957
2	Manhattan	New York	1628706	600.244	368500	22.83	59.13	71341	27544
3	Queens	Queens	2253858	93.310	41400	108.53	281.09	20767	8018
4	Staten Island	Richmond	476143	14.514	30500	58.37	151.18	8157	3150





Form dataset 1 we conclude the following:

- The highest Land area in Queens and the smallest in Manhattan
- The highest Population in Brooklyn, Queens and Manhattan.
- The highest Population Density in Manhattan, Brooklyn and The Bronx.
- Manhattan contributed in the highest GDP compared to all other Boroughs in New York City

The main insight of New York City data exploration is to select New York City Borough of Manhattan as the best location for the new Turkish restaurant (Manhattan has the best GDP and population density compared to all other Boroughs).

2-2 Manhattan Data Acquisition and Cleaning

In order to explore competitors in the selected borough in New York City we use the second dataset which was scraped from Wikipedia pages:

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

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

There are two lists in the second dataset: one of Michelin-starred restaurants 2020 and cuisine of New York City borough of Manhattan as given in the following tables

Second dataset

Name	Borough	
0	15 East	Manhattan
4	Agern	Manhattan
5	Al Fiori	Manhattan
7	Aldea	Manhattan
13	L'Appart	Manhattan
14	Aquavit	Manhattan
15	Aska	Brooklyn
16	L'Atelier de Joël Robuchon	Manhattan
17	Atera	Manhattan
18	Atomix	Manhattan
19	Aureole	Manhattan

Manhattan Michelin starred restaurants 2020

	Borough	Neighborhood	Cuisine
0	Manhattan	Chinatown	Chinese,Vietnamese
1	Manhattan	East Harlem	Puerto Rican, Mexican, Dominican, Chinese-Cuba...
2	Manhattan	East Village	Japanese, Korean, Indian, Ukrainian
3	Manhattan	Greenwich Village	Italian, Middle Eastern
4	Manhattan	Harlem	Italian, African-American, Latin American, We...
5	Manhattan	Koreatown	Korean
6	Manhattan	Nolita	Australian
7	Manhattan	Little Italy	Italian
8	Manhattan	Lower East Side	Puerto Rican, Jewish, Italian, Latin American
9	Manhattan	Murray Hill	Indian, Pakistani and Bangladeshi
10	Manhattan	Upper West Side, Manhattan	Jewish, Chinese-Latino
11	Manhattan	Washington Heights	Dominican, Puerto Rican, Mexican, Jewish
12	Manhattan	Upper East Side	German, Czech, Hungarian

cuisine of Manhattan

2-3 Manhattan Data Exploration

In order to segment New York City neighborhoods and explore them, we will essentially need a dataset that contains the 5 boroughs and the neighborhoods that exist in each borough as well as the latitude and longitude coordinates of each neighborhood.

Luckily, this dataset exists for free on the web. Feel free to try to find this dataset on your own, but here is the link to the dataset: https://geo.nyu.edu/catalog/nyu_2451_34572.

New York City 5 boroughs and their neighborhoods geographical coordinates and category data in the third dataset will be utilized as input for the Foursquare API, that will be leveraged to provision neighborhood venues information for selected borough (Manhattan). We will use the Foursquare API to explore neighborhoods in Manhattan. The below is Manhattan dataset obtained using Foursquare API data.

Third dataset

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Marble Hill	40.876551	-73.91066	Arturo's	40.874412	-73.910271	Pizza Place
1	Marble Hill	40.876551	-73.91066	Bikram Yoga	40.876844	-73.906204	Yoga Studio
2	Marble Hill	40.876551	-73.91066	Tibbett Diner	40.880404	-73.908937	Diner
3	Marble Hill	40.876551	-73.91066	Starbucks	40.877531	-73.905582	Coffee Shop
4	Marble Hill	40.876551	-73.91066	Dunkin'	40.877136	-73.906666	Donut Shop

Manhattan neighborhoods and venues geographical coordinates and category

3- Manhattan Neighborhoods Clustering

In order to select the best neighborhood in Manhattan for a new Turkish restaurant we will use the Foursquare API to explore neighborhoods in Manhattan.

we use the explore function to get the most common venue categories in each neighborhood.

Fourth dataset

The fourth dataset explore Manhattan neighborhoods along with the top 10 most common venues

	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	Park	Memorial Site	Food Court	Coffee Shop	Building	Gym	Shopping Mall	Burrito Place	Monument / Landmark	Movie Theater
1	Carnegie Hill	Gym / Fitness Center	Café	Bookstore	Pizza Place	Coffee Shop	Italian Restaurant	Gym	Gourmet Shop	Shipping Store	French Restaurant
2	Central Harlem	Chinese Restaurant	African Restaurant	American Restaurant	French Restaurant	Cafeteria	Bookstore	Boutique	Fried Chicken Joint	Spa	Food Truck
3	Chelsea	Ice Cream Shop	Theater	Cupcake Shop	American Restaurant	Speakeasy	Coffee Shop	New American Restaurant	Bar	Café	Nightclub
4	Chinatown	Chinese Restaurant	Hotpot Restaurant	Spa	Sandwich Place	Yoga Studio	Pizza Place	Boutique	Bubble Tea Shop	Spanish Restaurant	Salon / Barbershop

Manhattan neighborhoods top 10 most common venues

Cluster 1

	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
1	Chinatown	Chinese Restaurant	Hotpot Restaurant	Spa	Sandwich Place	Yoga Studio	Pizza Place	Boutique	Bubble Tea Shop	Spanish Restaurant	Salon / Barbershop
3	Inwood	Café	Deli / Bodega	Wine Bar	Park	Frozen Yogurt Shop	Restaurant	Bakery	Mexican Restaurant	Spanish Restaurant	Pet Store
11	Roosevelt Island	Park	Playground	School	Food & Drink Shop	Farmers Market	Cosmetics Shop	Bubble Tea Shop	Soccer Field	Bus Line	Dry Cleaner
21	Tribeca	Greek Restaurant	American Restaurant	Park	Men's Store	Yoga Studio	Hotel	Spa	Salad Place	Poke Place	Playground
26	Morningside Heights	American Restaurant	Bookstore	Park	Coffee Shop	Burger Joint	Sandwich Place	Mexican Restaurant	Greek Restaurant	Farmers Market	Café
28	Battery Park City	Park	Memorial Site	Food Court	Coffee Shop	Building	Gym	Shopping Mall	Burrito Place	Monument / Landmark	Movie Theater
32	Civic Center	French Restaurant	Gym / Fitness Center	Spa	Cocktail Bar	Park	Nail Salon	General Entertainment	Taco Place	Sushi Restaurant	Furniture / Home Store
35	Turtle Bay	Spa	Karaoke Bar	Coffee Shop	French Restaurant	Japanese Restaurant	Tennis Court	Gift Shop	Sushi Restaurant	Residential Building (Apartment / Condo)	Duty-free Shop
36	Tudor City	Park	Mexican Restaurant	Yoga Studio	Seafood Restaurant	Garden	Sushi Restaurant	Spanish Restaurant	Bridge	Burger Joint	Café
37	Stuyvesant Town	Park	Playground	Gym / Fitness Center	Skating Rink	Fountain	Bistro	Gas Station	Baseball Field	Farmers Market	Cocktail Bar
39	Hudson Yards	Hotel	American Restaurant	Gym / Fitness Center	Peruvian Restaurant	Coffee Shop	Café	Public Art	Residential Building (Apartment / Condo)	Furniture / Home Store	Supermarket

Cluster one is second biggest cluster with 11 of 39 neighbors in Manhattan. Up on closely examining these neighborhoods we can see that the most common venues are parks, restaurants, SPA, Gym, and coffee shops.

Cluster 2

Cluster two is the biggest cluster with 16 of 39 neighbors in Manhattan. Up on closely examining these neighborhoods we can see that the most common venues are restaurants and coffee shops

	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
2	Washington Heights	Café	Wine Shop	Park	Deli / Bodega	Tapas Restaurant	Indian Restaurant	Scenic Lookout	Frozen Yogurt Shop	Market	Breakfast Spot
4	Hamilton Heights	Coffee Shop	Yoga Studio	Pizza Place	Mexican Restaurant	Caribbean Restaurant	Cocktail Bar	Italian Restaurant	Pub	Park	Mediterranean Restaurant
6	Central Harlem	Chinese Restaurant	African Restaurant	American Restaurant	French Restaurant	Cafeteria	Bookstore	Boutique	Fried Chicken Joint	Spa	Food Truck
7	East Harlem	Mexican Restaurant	Latin American Restaurant	Thai Restaurant	Bakery	Sandwich Place	French Restaurant	Steakhouse	Cuban Restaurant	Café	Seafood Restaurant
10	Lenox Hill	Thai Restaurant	Burger Joint	Gym	Cycle Studio	Japanese Restaurant	French Restaurant	Chinese Restaurant	Liquor Store	Lingerie Store	Restaurant
12	Upper West Side	Italian Restaurant	American Restaurant	Bakery	Yoga Studio	Ramen Restaurant	Pub	Peruvian Restaurant	Nail Salon	Movie Theater	Middle Eastern Restaurant
15	Midtown	Hotel	Coffee Shop	Clothing Store	Tailor Shop	Historic Site	Spa	French Restaurant	Food Truck	Food Stand	Steakhouse
16	Murray Hill	Japanese Restaurant	Coffee Shop	Hotel	Burger Joint	Bagel Shop	Sushi Restaurant	Jewish Restaurant	Bar	Cocktail Bar	Ramen Restaurant
18	Greenwich Village	Italian Restaurant	Café	Sushi Restaurant	Clothing Store	French Restaurant	Yoga Studio	Jazz Club	Food Truck	Seafood Restaurant	New American Restaurant
19	East Village	Korean Restaurant	Wine Bar	Vietnamese Restaurant	Dessert Shop	Caribbean Restaurant	Speakeasy	Beer Store	Swiss Restaurant	Juice Bar	Moroccan Restaurant
20	Lower East Side	Chinese Restaurant	Café	Coffee Shop	Cocktail Bar	Art Gallery	Yoga Studio	Mediterranean Restaurant	French Restaurant	Mexican Restaurant	Filipino Restaurant
22	Little Italy	Sandwich Place	Coffee Shop	Wine Bar	Pizza Place	Ice Cream Shop	Bakery	French Restaurant	Salon / Barbershop	Salad Place	Chocolate Shop
23	Soho	Italian Restaurant	Boutique	Women's Store	Men's Store	Bakery	Mediterranean Restaurant	Electronics Store	Furniture / Home Store	Supermarket	Salon / Barbershop
24	West Village	Cocktail Bar	Bakery	Gourmet Shop	Italian Restaurant	Coffee Shop	Wine Bar	Chinese Restaurant	French Restaurant	Breakfast Spot	Candy Store
25	Manhattan Valley	Yoga Studio	Pizza Place	Mexican Restaurant	Coffee Shop	Bar	Vietnamese Restaurant	Thai Restaurant	Fried Chicken Joint	Ethiopian Restaurant	Latin American Restaurant
38	Flatiron	Coffee Shop	New American Restaurant	Japanese Restaurant	Yoga Studio	Café	Furniture / Home Store	Sushi Restaurant	Sports Club	Sporting Goods Shop	Miscellaneous Shop

Cluster 3

Cluster three has two neighbors, the most common venues are theaters and restaurants

	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
13	Lincoln Square	Theater	Indie Movie Theater	Concert Hall	Performing Arts Venue	Plaza	Gym	Cosmetics Shop	Park	Cycle Studio	Food Truck
14	Clinton	Theater	Gym / Fitness	Sandwich Place	Mediterranean Restaurant	Sporting Goods Shop	French Restaurant	Cocktail Bar	Coffee Shop	Lounge	Comedy Club

Cluster 4

Cluster four has only one neighbor, the most common venues are theaters and restaurants

	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
33	Midtown South	Korean Restaurant	Hotel	Scenic Lookout	Dessert Shop	Coffee Shop	Burger Joint	Building	Fried Chicken Joint	Steakhouse	Big Box Store

Cluster 5

Cluster five is the third biggest cluster with 10 of 39 neighbors in Manhattan. Up on closely examining these neighborhoods we can see that the most common venues are coffee shops, stores and shops

	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	Sandwich Place	Gym	Coffee Shop	Yoga Studio	Kids Store	Supplement Shop	Steakhouse	Miscellaneous Shop	Seafood Restaurant	Donut Shop
5	Manhattanville	Coffee Shop	Seafood Restaurant	Mexican Restaurant	Italian Restaurant	Indian Restaurant	Sushi Restaurant	Supermarket	Lounge	Boutique	Café
8	Upper East Side	French Restaurant	Hotel	Italian Restaurant	Gym / Fitness Center	Bakery	Burrito Place	Sculpture Garden	Sandwich Place	Bar	Salad Place
9	Yorkville	Italian Restaurant	Wine Shop	Deli / Bodega	Park	Coffee Shop	Hobby Shop	Gym	Burger Joint	Café	Monument / Landmark
17	Chelsea	Ice Cream Shop	Theater	Cupcake Shop	American Restaurant	Speakeasy	Coffee Shop	New American Restaurant	Bar	Café	Nightclub
27	Gramercy	Bagel Shop	Coffee Shop	Pizza Place	Yoga Studio	Bar	Sushi Restaurant	Spa	Cocktail Bar	Comedy Club	Convenience Store
29	Financial District	Coffee Shop	Event Space	Pizza Place	Gym / Fitness Center	Roof Deck	Mediterranean Restaurant	New American Restaurant	French Restaurant	Food Court	Steakhouse
30	Carnegie Hill	Gym / Fitness Center	Café	Bookstore	Pizza Place	Coffee Shop	Italian Restaurant	Gym	Gourmet Shop	Shipping Store	French Restaurant
31	Noho	Italian Restaurant	Wine Shop	Coffee Shop	Grocery Store	French Restaurant	Mexican Restaurant	Bookstore	Cocktail Bar	Sandwich Place	Gourmet Shop
34	Sutton Place	Gym / Fitness Center	Yoga Studio	Grocery Store	Italian Restaurant	Beer Garden	Gym	Mexican Restaurant	Liquor Store	Ice Cream Shop	Hotel

4- Discussion

This project aims to select the best location for a new Turkish restaurant in New York City. The selection criteria based on selecting one borough from 5 boroughs in New York City and then explore the neighborhoods of the selected borough to find the most 10 common Venues in each neighborhoods and finely clustering the neighborhoods use the k-means clustering algorithm. This select criteria manly based on the most common venues in each cluster where the category of venues in each cluster give a good indication of the type of business in this cluster. For a new Turkish restaurant after exploration of each cluster we found that clusters two, one and five respectively are the most suitable clusters for a new Turkish restaurant. The majority of venues in these clusters are restaurants, coffee shops and entertainment facilities (cinema, theater, shopping spa, malls, park, garden and playground).

5- Conclusion

In order to start a new Turkish restaurant in New York City we need to analyze restaurants business environment carefully. Our analysis based on selecting one borough from 5 boroughs in New York City and then explore the neighborhoods of the selected borough to find the category of the most common venues in these neighborhoods.

The insights derived from our analysis will give good understanding of restaurants business environment in New York City and this will have a major effect on reduce risk of investment and increase Return of Investment (RoI) in this highly competitive market.

The man insights derived from our analysis are:

- New York City Borough of Manhattan is the best location for the new Turkish restaurant (Manhattan has the best GDP and population density compared to all other Boroughs)
- clusters two, one and five respectively are the most suitable clusters for a new Turkish restaurant. The majority of venues in these clusters are restaurants, coffee shops and entertainment facilities (cinema, theater, coffee shops, shops, spa, malls, park, garden and playground).