Introduction & Business Problem:

The City of New York, is the most populous city in the United States. It is diverse and is the financial capital of USA. It is multicultural. It provides lot of business opportunities and business friendly environment. It has attracted many different players into the market. It is a global hub of business and commerce. The city is a major 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. This also means that the market is highly competitive. As it is highly developed city so cost of doing business is also one of the highest. Thus, any new business venture or expansion needs to be analyzed carefully. The insights derived from analysis will give good understanding of the business environment which help in strategically targeting the market. This will help in reduction of risk. And the Return on Investment will be reasonable.

Business Problem

It's something in the water or at least that's what they'll have you believe. One of the oldest theories about New York City pizza is that the city's tap water is responsible for its superior taste. Pizza may have been invented in Italy, but it was perfected in New York City. Thus, my client would like to start his own legacy in the Pizza making business. Pizza places are unlimited in number in Manhattan but the demand is very high that the number of places is just not enough. Various factors need to be weighed in while choosing the best place to put up a pizza store. The objective of this project will be to analyse and choose the best neighbourhood to start a pizza place in NYC and specifically in Manhattan.

Data

Neighborhood has a total of 5 boroughs and 306 neighborhoods. In order to segment the 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. This dataset exists for free on the web. Link to the dataset is: http://cocl.us/new_york_dataset

	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

New York city geographical coordinates data will be utilized as input for the Foursquare API, that will be leveraged to provision venues information for each neighborhood. We will use the Foursquare API to explore neighborhoods in New York City. The below is image of the Foursquare API data. In addition, Pizza category Id 4bf58dd8d48988d1ca941735 is used for retrieving data from Foursquare API.

Methodology

In this project, I will use the basic methodology as taught in Week 3 lab.

	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

Above, I have done conversion of addresses into their equivalent latitude and longitude values. Then we will use the Foursquare API to explore neighborhoods in Manhattan, New York. After that, we will explore a function to get pizza restaurant categories in each neighborhood.

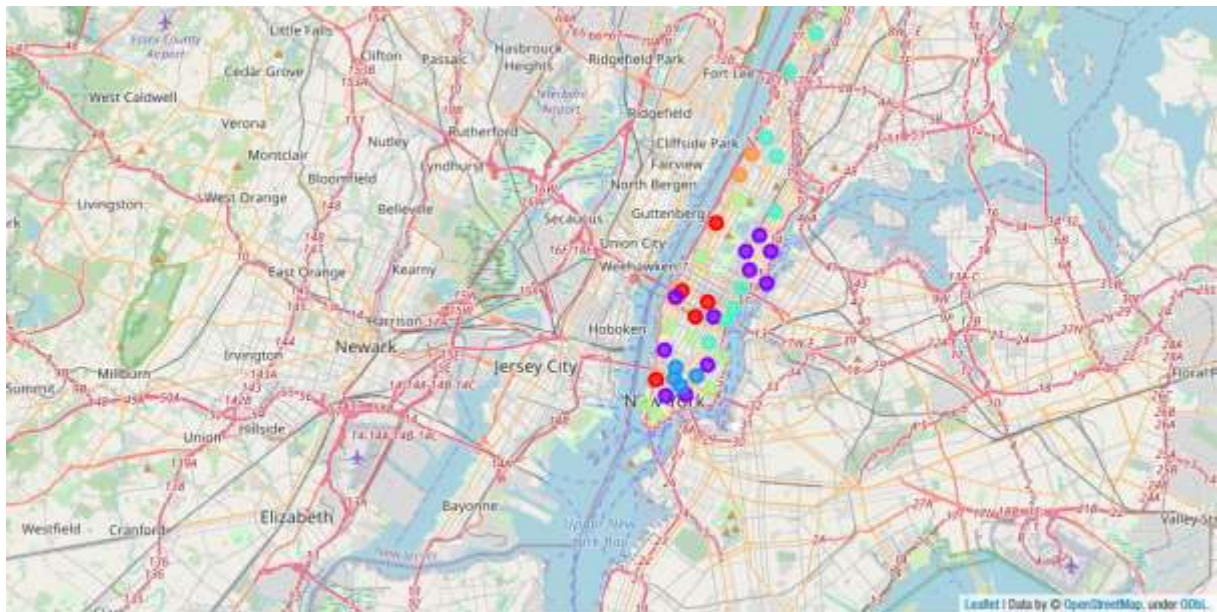|	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Marble Hill	40.876551	-73.91066	Artaro's	40.874412	-73.910271	Pizza Place
1	Marble Hill	40.876551	-73.91066	Cafeccino Bakery	40.880209	-73.906884	Bagel Shop
2	Marble Hill	40.876551	-73.91066	Supreme Pizza & Pasta	40.881001	-73.908966	Pizza Place
3	Marble Hill	40.876551	-73.91066	Sam's Pizza	40.879435	-73.905859	Pizza Place
4	Marble Hill	40.876551	-73.91066	Broadway Pizza & Pasta	40.878822	-73.904494	Pizza Place

```
In [20]: newyork_pizza.shape
```
```
Out[20]: (1656, 7)
```


Map of Pizza places in Manhattan.

Then we will use this feature to group the neighborhoods into clusters K-means clustering algorithm will be use to complete this task. And also, the Folium library to visualize the neighborhoods in Manhattan and its emerging clusters.

	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	Pizza Place	Italian Restaurant	Kosher Restaurant	Gourmet Shop	Turkish Restaurant	Deli / Bodega	Arcade	Bagel Shop	Bakery	Bar
1	Carnegie Hill	Pizza Place	Italian Restaurant	Turkish Restaurant	Diner	Arcade	Bagel Shop	Bakery	Bar	Boat or Ferry	Burger Joint
2	Central Harlem	Pizza Place	Deli / Bodega	Turkish Restaurant	Diner	Arcade	Bagel Shop	Bakery	Bar	Boat or Ferry	Burger Joint
3	Chelsea	Pizza Place	Deli / Bodega	Mexican Restaurant	Italian Restaurant	Café	Turkish Restaurant	Arcade	Bagel Shop	Bakery	Bar
4	Chinatown	Pizza Place	Italian Restaurant	Turkish Restaurant	Diner	Arcade	Bagel Shop	Bakery	Bar	Boat or Ferry	Burger Joint

Map after clustering.

Results

Using K-mean to cluster data area and obtain cluster with less number of pizza places.

Cluster 0


```
In [49]: manhattan_merged.loc[manhattan_merged['Cluster Labels'] == 0, manhattan_merged.columns[[1] + list(range(5, manhattan_merged.shape
```

Out[49]:

	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
12	Upper West Side	Pizza Place	Italian Restaurant	Turkish Restaurant	Bakery	American Restaurant	Kosher Restaurant	Cocktail Bar	Arcade	Bagel Shop	Sandwich Place
14	Clinton	Pizza Place	Italian Restaurant	Bar	Fast Food Restaurant	Sandwich Place	Deli / Bodega	Arcade	Bagel Shop	Bakery	Boat or Ferry
15	Midtown	Pizza Place	Italian Restaurant	Sandwich Place	American Restaurant	Deli / Bodega	Greek Restaurant	Arcade	Bagel Shop	Bakery	Bar
21	Tribeca	Pizza Place	Italian Restaurant	Bagel Shop	Bakery	Food Truck	Sports Bar	Turkish Restaurant	Deli / Bodega	Arcade	Bar
33	Midtown South	Pizza Place	Italian Restaurant	American Restaurant	Mexican Restaurant	Greek Restaurant	Deli / Bodega	Arcade	Bagel Shop	Bakery	Bar

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	Pizza Place	Italian Restaurant	Turkish Restaurant	Diner	Arcade	Bagel Shop	Bakery	Bar	Boat or Ferry	Burger Joint
8	Upper East Side	Pizza Place	Italian Restaurant	Turkish Restaurant	Diner	Arcade	Bagel Shop	Bakery	Bar	Boat or Ferry	Burger Joint
9	Yorkville	Pizza Place	Italian Restaurant	Turkish Restaurant	Diner	Arcade	Bagel Shop	Bakery	Bar	Boat or Ferry	Burger Joint
10	Lenox Hill	Pizza Place	Italian Restaurant	Turkish Restaurant	Diner	Arcade	Bagel Shop	Bakery	Bar	Boat or Ferry	Burger Joint
11	Roosevelt Island	Pizza Place	Italian Restaurant	Turkish Restaurant	Diner	Arcade	Bagel Shop	Bakery	Bar	Boat or Ferry	Burger Joint
16	Murray Hill	Pizza Place	Italian Restaurant	American Restaurant	Sandwich Place	Deli / Bodega	Arcade	Bagel Shop	Bakery	Bar	Boat or Ferry
19	East Village	Pizza Place	Italian Restaurant	Bar	Cocktail Bar	Turkish Restaurant	Diner	Arcade	Bagel Shop	Bakery	Boat or Ferry
24	West Village	Pizza Place	Italian Restaurant	Gluten-free Restaurant	Café	Turkish Restaurant	Deli / Bodega	Arcade	Bagel Shop	Bakery	Bar
30	Carnegie Hill	Pizza Place	Italian Restaurant	Turkish Restaurant	Diner	Arcade	Bagel Shop	Bakery	Bar	Boat or Ferry	Burger Joint
32	Civic Center	Pizza Place	Italian Restaurant	Bagel Shop	Bakery	Turkish Restaurant	Diner	Arcade	Bar	Boat or Ferry	Burger Joint
39	Hudson Yards	Pizza Place	Italian Restaurant	Bar	Deli / Bodega	Turkish Restaurant	Diner	Arcade	Bagel Shop	Bakery	Boat or Ferry

Cluster 2

	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
18	Greenwich Village	Pizza Place	Italian Restaurant	Sandwich Place	Bagel Shop	Bakery	Turkish Restaurant	Deli / Bodega	Arcade	Bar	Boat or Ferry
22	Little Italy	Pizza Place	Italian Restaurant	Bagel Shop	Bakery	Turkish Restaurant	Diner	Arcade	Bar	Boat or Ferry	Burger Joint
23	Soho	Pizza Place	Italian Restaurant	Bagel Shop	Bakery	Turkish Restaurant	Diner	Arcade	Bar	Boat or Ferry	Burger Joint
31	Noho	Pizza Place	Italian Restaurant	Cocktail Bar	Turkish Restaurant	Diner	Arcade	Bagel Shop	Bakery	Bar	Boat or Ferry

Cluster 3

	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	Pizza Place	Café	Turkish Restaurant	Diner	Arcade	Bagel Shop	Bakery	Bar	Boat or Ferry	Burger Joint
3	Inwood	Pizza Place	Turkish Restaurant	Diner	Arcade	Bagel Shop	Bakery	Bar	Boat or Ferry	Burger Joint	Café
4	Hamilton Heights	Pizza Place	Deli / Bodega	Turkish Restaurant	Diner	Arcade	Bagel Shop	Bakery	Bar	Boat or Ferry	Burger Joint
6	Central Harlem	Pizza Place	Deli / Bodega	Turkish Restaurant	Diner	Arcade	Bagel Shop	Bakery	Bar	Boat or Ferry	Burger Joint
7	East Harlem	Pizza Place	American Restaurant	Deli / Bodega	Diner	Arcade	Bagel Shop	Bakery	Bar	Boat or Ferry	Burger Joint
27	Gramercy	Pizza Place	Italian Restaurant	Bar	Turkish Restaurant	Diner	Arcade	Bagel Shop	Bakery	Boat or Ferry	Burger Joint
34	Sutton Place	Pizza Place	Sandwich Place	Italian Restaurant	Turkish Restaurant	Deli / Bodega	Arcade	Bagel Shop	Bakery	Bar	Boat or Ferry
35	Turtle Bay	Pizza Place	Sandwich Place	Italian Restaurant	Turkish Restaurant	Deli / Bodega	Arcade	Bagel Shop	Bakery	Bar	Boat or Ferry
36	Tudor City	Pizza Place	Italian Restaurant	Turkish Restaurant	Diner	Arcade	Bagel Shop	Bakery	Bar	Boat or Ferry	Burger Joint

Cluster 4

	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	Pizza Place	Bagel Shop	Turkish Restaurant	Diner	Arcade	Bakery	Bar	Boat or Ferry	Burger Joint	Café
13	Lincoln Square	Pizza Place	Deli / Bodega	Bakery	Italian Restaurant	Burger Joint	Turkish Restaurant	Diner	Arcade	Bagel Shop	Bar
17	Chelsea	Pizza Place	Deli / Bodega	Mexican Restaurant	Italian Restaurant	Café	Turkish Restaurant	Arcade	Bagel Shop	Bakery	Bar
20	Lower East Side	Pizza Place	Italian Restaurant	Cocktail Bar	Turkish Restaurant	Diner	Arcade	Bagel Shop	Bakery	Bar	Boat or Ferry
25	Manhattan Valley	Pizza Place	Deli / Bodega	Italian Restaurant	Fried Chicken Joint	Turkish Restaurant	Arcade	Bagel Shop	Bakery	Bar	Boat or Ferry
28	Battery Park City	Pizza Place	Italian Restaurant	Kosher Restaurant	Gourmet Shop	Turkish Restaurant	Deli / Bodega	Arcade	Bagel Shop	Bakery	Bar
29	Financial District	Pizza Place	Italian Restaurant	Deli / Bodega	Bar	Gourmet Shop	Food Truck	Turkish Restaurant	Arcade	Bagel Shop	Bakery
37	Stuyvesant Town	Pizza Place	Italian Restaurant	Bar	Boat or Ferry	Cocktail Bar	Turkish Restaurant	Diner	Arcade	Bagel Shop	Bakery
38	Flatiron	Pizza Place	Italian Restaurant	Sandwich Place	Mexican Restaurant	Café	Turkish Restaurant	Deli / Bodega	Arcade	Bagel Shop	Bakery

Cluster 5

	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
5	Manhattanville	Pizza Place	Diner	Arcade	Italian Restaurant	Turkish Restaurant	Bagel Shop	Bakery	Bar	Boat or Ferry	Burger Joint
26	Morningside Heights	Pizza Place	Italian Restaurant	Diner	Arcade	Deli / Bodega	Turkish Restaurant	Bagel Shop	Bakery	Bar	Boat or Ferry

Discussion

This analysis is performed on limited data. This may be right or may be wrong. But if good amount of data is available there is scope to come up with better results.
Based on data frame analysis above Cluster 2 – Greenwich Village, Noho, Soho, Little Italy and Cluster 5 - Morningside Heights, Manhattanville areas are the best places to open a new Pizza place business.

There is high competition in Clusters 1,3,4 - China town, Washington Heights and Marble Hill, so it is very risky to open business in these areas.
The analysis can be made better by obtaining people demographics of people living and another major problem with pizzas faced is the delivery timings. So the delivery routes and coverage will also play a major factor. Finally, Four Square proved to be a good source of data.

Conclusion

Although all of the goals of this project were met there is definitely room for further improvement and development. The project could easily be developed into a fully-fledged application that could support the opening of a business in an unknown location.

The same method can be used to find for any type of restaurant or any service of that type.