

# Capstone Data Science Project

## The Battle of the Neighborhoods!



## New York

## Food Industry!

*-As part of completing the IBM Applied Data  
Science Capstone Course*

## Contents

1	Introduction .....	3
2	Business Case .....	3
3	Data.....	3
4	Methodology .....	4
5	Results .....	5
5.1	K means Cluster Results .....	5
5.1.1	Cluster Tables .....	5
5.1.2	Map Showing the locations of each cluster.....	8
5.1.3	Intermediate Discussion .....	8
5.2	Frequency Table Results .....	9
5.2.1	Upper Class areas in New York .....	9
5.2.2	Middle Class areas in New York.....	10
5.2.3	Lower Class areas in New York.....	10
6	Discussion.....	11
6.1	Observation #1 .....	11
6.2	Observation #2 .....	11
6.3	Observation #3 .....	11
6.4	Observation #4 .....	11
6.5	Observation #5 .....	11
6.6	Recommendations .....	12
7	Conclusion.....	12

# 1 Introduction

Businessmen are always attempting to expand their wealth by either growing their current companies or expanding their current portfolio in order to retrieve more streams of income. Ideally this is always informed through a complex decision matrix on possible expansion strategies they wish to pursue. Business ideas are brainstormed, the top ideas are selected and business plans are written out on a high level. The top business plans are further developed and eliminated as the level of detail increases and finally out of this entire process, an idea becomes a project which is pursued. Through many iterations and blending of business ideas, the final idea may not have been an idea from the iteration process but something that was constructed during the iteration stage. One of the important elements in refining a business idea is always to understand a market and potential customers. There are many traits that characterise potential customers such as lifestyle, income bracket, culture, beliefs, location, etc. By identifying the potential customers and characterising them, a potential investor is able to decide where they wish to pursue their next endeavour. This is one of the many ways Data Scientists are able to use data to understand communities and better suggest products/services that will indeed be successful due to the fact that it makes lives easier for people. The most successful businesses are the ones that improve the quality of life for people and bring the most value to their respective customers.

## 2 Business Case

Our client owns businesses in the food industry across the world and have finally grown enough to infiltrate business into the New York area. The client hired our company to investigate various business opportunities in the New York area and provide them with our suggestions on the various business ventures they may partake in. As this is the first time that our client is looking into New York City, they require us to report on the various business opportunities (with specific interest in the food industry) we find and provide recommendations on what their strategy should be. As a Data Scientist within the company, we are required to understand the various areas within New York by understanding the current businesses that exist within New York and provide the strategy team with more locational information according to the neighbourhoods so that they can better support their business cases that they wish to recommend to the client.

## 3 Data

Our data will make use of the FourSquare API and will gather the various types of food businesses that exist within the various neighborhoods of New York. These businesses will then be categorised according to the types of businesses that they are such as the various types of restaurants. The locational data of New York will be retrieved through a geojson file. An article off the internet will also be used to scrape Median Household Income (MHI) from respective neighborhoods from the internet as it will be interesting to observe the types of business that exist from neighbourhood trends due to income brackets. This will give us a better understanding of why some food businesses exist within neighbourhoods and better understand the picture of how businesses operate within New York. By observing the output of the study, the relationships between the neighbourhoods, income, and business will be seen easier which will help our client decide on the various businesses that need to be pursued according to the neighbourhood requirements as well as the various trends that exist between similar neighborhoods.

## 4 Methodology

Firstly, I would like to state my hypothesis. Due to no knowledge on New York City this may be difficult to be very specific on what it is we are looking for since we are just seeking understanding more than proving something. With this in mind, my hypothesis is:

1. The food industry has a market in every neighborhood regardless of MHI as food is a basic need for humans
2. There will be differences between the neighborhoods in terms of the types of food businesses that exist as poorer neighborhoods will cover mostly necessities whereas wealthier neighborhoods can afford a higher lifestyle standard
3. Neighborhoods that do not share traits with their relevant MHI counterparts is a possible business opportunity as there may be market share that has not been utilised yet that can place it in similar standing to neighborhoods in its MHI category.

To go about the data analysis, I will request food business data from the foursquare API for each neighborhood within a radius of 500m from the neighborhood's centre with a limit of 150 businesses. The distance and limit has been changed a few times with the results being saved to a csv file so that previously retrieved data can be reused without using the API again due to the limit on the foursquare account. The data is received in the form of a json file and the data extracted with a defined function to retrieve the business name, category and locational data.

To analyse the neighborhood data, we wish to see which are the top 10 most common places within each neighborhood. To do this, we use the one hot encoding method, group all the neighborhoods with each category of food businesses being stated as an average of the total amount of food businesses (normalising the data). This is done so that we can sort the data and allocate the top ten most common venues within each neighborhood. This table is then fed into a K means algorithm to see which types of neighborhoods are similar according to the normalised quantity of restaurants they find in the area. The number of clusters in the machine learning algorithm is an iterative process by changing the number of clusters required and observing each results before deciding on an optimum. The MHI data will be added back to the clustered data and the results will be observed with hopefully differences being noted between various MHI brackets.

For further analysis to paint a deeper picture, the data will be manually split according to MHI with the three categories being Middle-Upper class(referred to as upper class for the rest of this report) Middle-Middle class (referred to as the middle class for the rest of this report) and Middle-lower class(referred to lower class for the rest of this report). The tallies of the top ten most common restaurants will be added up for each MHI bracket and the results will be observed and visualised.

Finally, the results will be investigated, and observations will be made that could aid other departments in making a suggestion for the client.

## 5 Results

The following section will look at the results following the data analysis.

### 5.1 K means Cluster Results

#### 5.1.1 Cluster Tables

For the sake of neatly displaying results, the results shown are only the top 5 most common food businesses per neighborhood but the analysis did use the top ten businesses in each area which is shown on the github jupyter notebook which can be found at:

[https://github.com/MatthewVT/Coursera\\_Capstone/blob/master/Week%205%20PG%20Assignment/Battle%20of%20the%20New%20York%20Neighbourhoods-%20Food%20Business%20Case.ipynb](https://github.com/MatthewVT/Coursera_Capstone/blob/master/Week%205%20PG%20Assignment/Battle%20of%20the%20New%20York%20Neighbourhoods-%20Food%20Business%20Case.ipynb)

##### 5.1.1.1 Cluster 1

Neighborhood	Median Household Income	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
Tribeca	\$ 186 416.00	American Restaurant	Italian Restaurant	Café	Deli / Bodega	Greek Restaurant
Carnegie Hill	\$ 171 603.00	Pizza Place	Café	Bakery	Italian Restaurant	Sushi Restaurant
Financial District	\$ 158 605.00	Sandwich Place	Salad Place	Italian Restaurant	Pizza Place	American Restaurant
Upper East Side	\$ 131 996.00	Italian Restaurant	American Restaurant	Pizza Place	French Restaurant	Bakery
Cobble Hill	\$ 131 817.00	Deli / Bodega	Thai Restaurant	Pizza Place	Bakery	Italian Restaurant
Soho	\$ 126 708.00	Italian Restaurant	Café	Mediterranean Restaurant	Sandwich Place	Seafood Restaurant
Greenwich Village	\$ 121 594.00	Italian Restaurant	Café	Sushi Restaurant	Vegetarian / Vegan Restaurant	Pizza Place
Upper West Side	\$ 121 502.00	Italian Restaurant	Indian Restaurant	Pizza Place	Café	Bakery
Murray Hill	\$ 117 763.00	Korean Restaurant	Japanese Restaurant	Deli / Bodega	Sandwich Place	Mediterranean Restaurant
Murray Hill	\$ 117 763.00	Korean Restaurant	Japanese Restaurant	Deli / Bodega	Sandwich Place	Mediterranean Restaurant
West Village	\$ 117 237.00	Italian Restaurant	New American Restaurant	American Restaurant	Seafood Restaurant	Café
Midtown	\$ 116 728.00	American Restaurant	Food Truck	Sushi Restaurant	Japanese Restaurant	Sandwich Place
Carroll Gardens	\$ 115 860.00	Italian Restaurant	Deli / Bodega	Pizza Place	Bakery	French Restaurant
Park Slope	\$ 115 042.00	Pizza Place	American Restaurant	Burger Joint	Chinese Restaurant	Italian Restaurant
Chelsea	\$ 112 319.00	French Restaurant	Italian Restaurant	Café	Bakery	American Restaurant
Chelsea	\$ 112 319.00	French Restaurant	Italian Restaurant	Café	Bakery	American Restaurant
Gramercy	\$ 111 522.00	Italian Restaurant	Bagel Shop	Pizza Place	Mexican Restaurant	Diner
Downtown	\$ 110 356.00	Pizza Place	Chinese Restaurant	Sandwich Place	American Restaurant	Food Truck
Woodrow	\$ 104 778.00	Donut Shop	Chinese Restaurant	Diner	Sushi Restaurant	Bakery
Boerum Hill	\$ 104 198.00	Deli / Bodega	Sandwich Place	Pizza Place	Food Truck	Seafood Restaurant
Huguenot	\$ 100 754.00	Deli / Bodega	Italian Restaurant	Donut Shop	Sandwich Place	Asian Restaurant
Great Kills	\$ 92 004.00	Pizza Place	Italian Restaurant	Bagel Shop	Mexican Restaurant	Spanish Restaurant

<b>Little Italy</b>	\$ 91 061.00	Chinese Restaurant	Italian Restaurant	Café	Vietnamese Restaurant	Bakery
<b>Clinton</b>	\$ 90 686.00	Italian Restaurant	Deli / Bodega	American Restaurant	Restaurant	Sandwich Place
<b>City Island</b>	\$ 84 445.00	Deli / Bodega	Seafood Restaurant	Spanish Restaurant	Diner	Pizza Place
<b>Morningside Heights</b>	\$ 77 589.00	Deli / Bodega	Pizza Place	Food Truck	American Restaurant	Café
<b>Forest Hills</b>	\$ 73 556.00	Deli / Bodega	Food Truck	Asian Restaurant	Thai Restaurant	Pizza Place
<b>Auburndale</b>	\$ 73 129.00	Deli / Bodega	American Restaurant	Italian Restaurant	Diner	Fast Food Restaurant
<b>East Village</b>	\$ 70 370.00	Pizza Place	Japanese Restaurant	Mexican Restaurant	Vietnamese Restaurant	Italian Restaurant
<b>Chinatown</b>	\$ 69 168.00	Chinese Restaurant	Bakery	Vietnamese Restaurant	Dim Sum Restaurant	Malay Restaurant
<b>Dyker Heights</b>	\$ 64 341.00	Italian Restaurant	Hot Dog Joint	Food Truck	Fast Food Restaurant	Food
<b>Bay Ridge</b>	\$ 64 284.00	Pizza Place	Italian Restaurant	Chinese Restaurant	Mediterranean Restaurant	Bagel Shop
<b>Steinway</b>	\$ 62 031.00	Deli / Bodega	Italian Restaurant	Café	Sushi Restaurant	Food Truck
<b>Country Club</b>	\$ 61 099.00	Sandwich Place	Italian Restaurant	Chinese Restaurant	Wings Joint	Empanada Restaurant
<b>Jackson Heights</b>	\$ 52 815.00	Latin American Restaurant	South American Restaurant	Peruvian Restaurant	Bakery	Mexican Restaurant
<b>Flushing</b>	\$ 47 960.00	Korean Restaurant	Chinese Restaurant	Bakery	Hotpot Restaurant	Deli / Bodega
<b>Yorkville</b>	\$ 46 148.00	Italian Restaurant	Pizza Place	Deli / Bodega	Sushi Restaurant	Sandwich Place

The mean MHI for cluster 1 is \$100 745.03 with a total of 37 neighborhoods.

### 5.1.1.2 Cluster 2

Neighborhood	Median Household Income	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
<b>Whitestone</b>	\$ 86 250.00	Deli / Bodega	Wings Joint	Food Court	Diner	Donut Shop
<b>South Beach</b>	\$ 80 005.00	Deli / Bodega	Wings Joint	Food Court	Diner	Donut Shop

The mean MHI for cluster 2 is \$83 127.50 with a total of 2 neighborhoods.

### 5.1.1.3 Cluster 3

Neighborhood	Median Household Income	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
<b>Spuyten Duyvil</b>	\$92 579.00	Thai Restaurant	Asian Restaurant	Fast Food Restaurant	Pizza Place	Food
<b>Rossville</b>	\$86 343.00	Pizza Place	Bagel Shop	BBQ Joint	Deli / Bodega	American Restaurant
<b>Midland Beach</b>	\$82 585.00	Deli / Bodega	Pizza Place	Restaurant	Chinese Restaurant	Café
<b>Rosebank</b>	\$82 109.00	Pizza Place	Mexican Restaurant	Deli / Bodega	Italian Restaurant	Bakery
<b>Middle Village</b>	\$76 088.00	Chinese Restaurant	Bakery	Deli / Bodega	Spanish Restaurant	South American Restaurant
<b>Queens Village</b>	\$73 128.00	Mexican Restaurant	Food	Bakery	Restaurant	Diner
<b>Utopia</b>	\$71 438.00	Deli / Bodega	Afghan Restaurant	Bakery	Donut Shop	Pizza Place
<b>Canarsie</b>	\$69 844.00	Food	Deli / Bodega	Asian Restaurant	Caribbean Restaurant	Chinese Restaurant
<b>Eastchester</b>	\$69 654.00	Caribbean Restaurant	Deli / Bodega	Diner	Pizza Place	Seafood Restaurant
<b>Baychester</b>	\$69 222.00	Donut Shop	Mexican Restaurant	Pizza Place	Fast Food Restaurant	Fried Chicken Joint
<b>Glendale</b>	\$69 037.00	Pizza Place	Deli / Bodega	Chinese Restaurant	Wings Joint	Diner
<b>Springfield Gardens</b>	\$68 252.00	Donut Shop	Fried Chicken Joint	Sandwich Place	Chinese Restaurant	Fast Food Restaurant
<b>New Brighton</b>	\$68 125.00	Deli / Bodega	Café	Chinese Restaurant	Bakery	Ethiopian Restaurant
<b>Maspeth</b>	\$61 497.00	Chinese Restaurant	Deli / Bodega	Pizza Place	Diner	Taco Place

<b>Clifton</b>	\$60 148.00	Mexican Restaurant	Pizza Place	Deli / Bodega	American Restaurant	Eastern European Restaurant
<b>Ridgewood</b>	\$58 599.00	Deli / Bodega	Pizza Place	Italian Restaurant	Bakery	Restaurant
<b>College Point</b>	\$57 743.00	Pizza Place	Deli / Bodega	Bakery	Latin American Restaurant	Asian Restaurant
<b>Morris Park</b>	\$57 355.00	Pizza Place	Deli / Bodega	Bakery	Food Truck	Burger Joint
<b>Queensboro Hill</b>	\$57 190.00	Chinese Restaurant	Asian Restaurant	Pizza Place	Cantonese Restaurant	Fast Food Restaurant
<b>Port Richmond</b>	\$55 390.00	Mexican Restaurant	Pizza Place	Deli / Bodega	Donut Shop	Food Truck
<b>Woodside</b>	\$53 882.00	Bakery	Latin American Restaurant	Deli / Bodega	Thai Restaurant	Pizza Place
<b>Williamsburg</b>	\$53 643.00	Deli / Bodega	Latin American Restaurant	Pizza Place	Breakfast Spot	Bagel Shop
<b>Bensonhurst</b>	\$53 204.00	Chinese Restaurant	Pizza Place	Italian Restaurant	Asian Restaurant	Sushi Restaurant
<b>Lower East Side</b>	\$49 246.00	Pizza Place	Chinese Restaurant	Café	Bakery	Deli / Bodega
<b>Corona</b>	\$49 118.00	Mexican Restaurant	Bakery	Italian Restaurant	Latin American Restaurant	Restaurant
<b>Parkchester</b>	\$48 567.00	Pizza Place	Deli / Bodega	Chinese Restaurant	American Restaurant	Asian Restaurant
<b>Inwood</b>	\$47 582.00	Pizza Place	Restaurant	Café	Mexican Restaurant	Spanish Restaurant
<b>Washington Heights</b>	\$46 288.00	Deli / Bodega	Pizza Place	Mexican Restaurant	Chinese Restaurant	Bakery
<b>Bushwick</b>	\$45 238.00	Deli / Bodega	Mexican Restaurant	Pizza Place	Sandwich Place	Chinese Restaurant
<b>Borough Park</b>	\$44 528.00	Deli / Bodega	Pizza Place	Fast Food Restaurant	Bakery	Café
<b>Hamilton Heights</b>	\$43 433.00	Deli / Bodega	Pizza Place	Mexican Restaurant	Café	Sandwich Place
<b>Brownsville</b>	\$43 390.00	Restaurant	Chinese Restaurant	Fried Chicken Joint	Pizza Place	Spanish Restaurant
<b>Sunset Park</b>	\$43 202.00	Mexican Restaurant	Deli / Bodega	Bakery	Pizza Place	Latin American Restaurant
<b>Bedford Park</b>	\$38 675.00	Deli / Bodega	Chinese Restaurant	Pizza Place	Mexican Restaurant	Diner
<b>Soundview</b>	\$35 677.00	Chinese Restaurant	Food	Breakfast Spot	Fried Chicken Joint	Latin American Restaurant
<b>East Harlem</b>	\$32 744.00	Mexican Restaurant	Pizza Place	Bakery	Deli / Bodega	Latin American Restaurant
<b>University Heights</b>	\$32 542.00	Pizza Place	Chinese Restaurant	Fried Chicken Joint	Burger Joint	Bakery
<b>Fordham</b>	\$29 604.00	Pizza Place	Spanish Restaurant	Donut Shop	Chinese Restaurant	Fast Food Restaurant
<b>Hunts Point</b>	\$26 421.00	Food	Spanish Restaurant	BBQ Joint	Pizza Place	Café
<b>Morris Heights</b>	\$26 278.00	Spanish Restaurant	Deli / Bodega	Chinese Restaurant	Latin American Restaurant	Pizza Place
<b>Mott Haven</b>	\$24 839.00	Spanish Restaurant	Donut Shop	Pizza Place	Bakery	Peruvian Restaurant

The mean MHI for cluster 3 is \$54 986.02 with a total of 41 neighborhoods.

#### 5.1.1.4 Cluster 4

Neighborhood	Median Household Income	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
<b>Riverdale</b>	\$ 86 341.00	Food Truck	Wings Joint	Food Court	Diner	Donut Shop

The mean MHI for cluster 4 is \$86 341.00 with a total of 1 neighborhood.

#### 5.1.1.5 Cluster 5

Neighborhood	Median Household Income	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
<b>Laurelton</b>	\$ 85 256.00	Caribbean Restaurant	Deli / Bodega	Wings Joint	Diner	Donut Shop

The mean MHI for cluster 5 is \$85 256.00 with a total of 1 neighborhood.



### 5.1.1.6 Cluster 6

Neighborhood	Median Household Income	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
Howland Hook	\$ 64 184.00	Italian Restaurant	Wings Joint	Food Court	Diner	Donut Shop

The mean MHI for cluster 6 is \$64 184.00 with a total of 1 neighborhood.

### 5.1.2 Map Showing the locations of each cluster

This map shows all the neighbourhoods that were clustered into one of 6 clusters according to the quantity of the food businesses in each area.

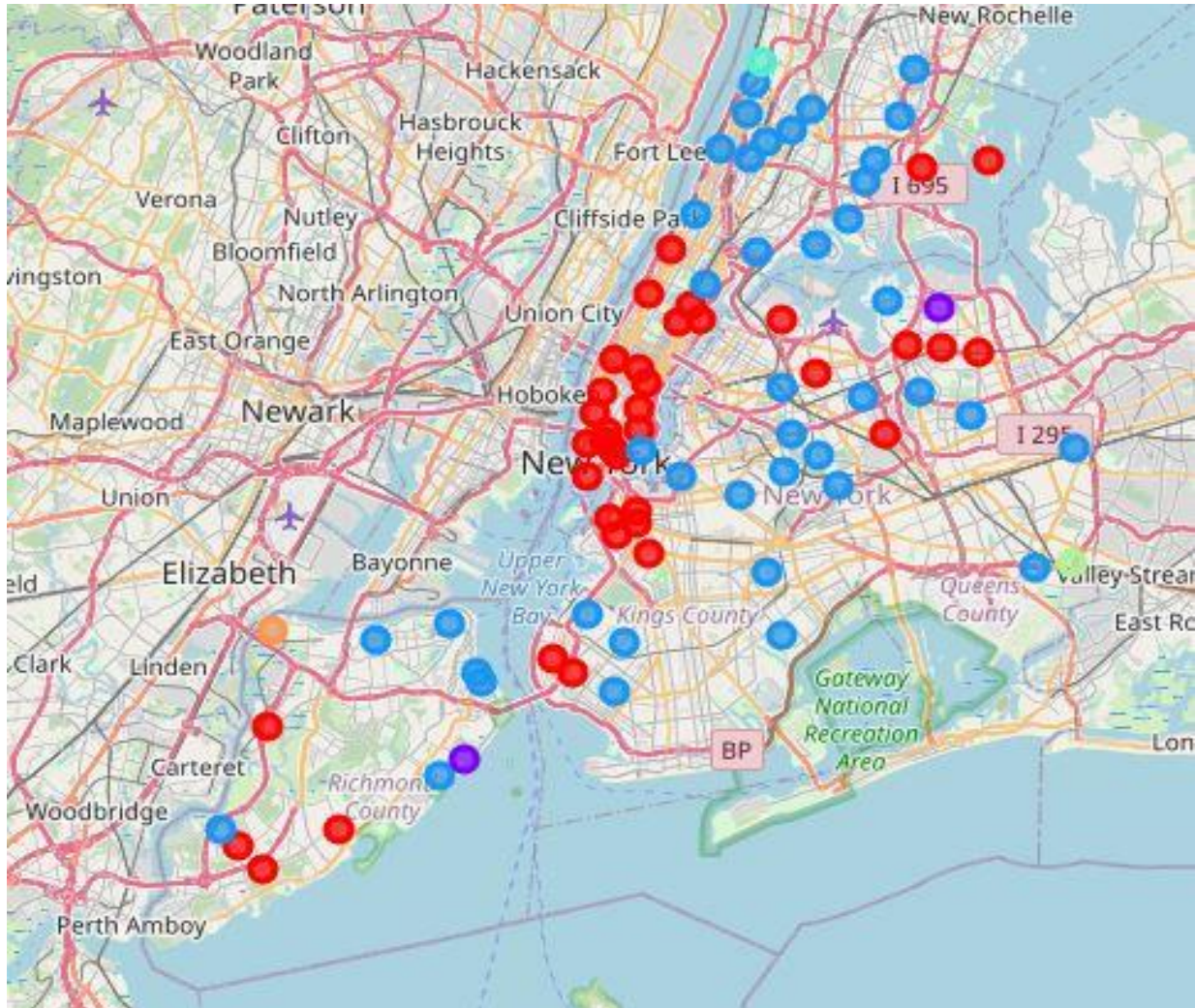


Figure 1: Folium map indicating neighborhood clusters

### 5.1.3 Intermediate Discussion

Looking at the clusters, there appears to be a division between the Upper class and the rest of the data set with a few outliers. The clusters 1, 3, 4 and 5 appear to be very similar to each other if you consider that there are about 7 similarities between them in their most common venues but due to the requirement of having 6 clusters the computer may have separated them to suit this requirement. The final number of clusters were chosen at 6. This was an iterative process of selecting clusters from 1 to 15 and observing the results. Using 6 clusters was found to be the most helpful in noticing somewhat a split between Upper class and the rest MHI brackets. Unfortunately, this does not really tell a convincing story in terms of why that is or how to suggest a businessman to suggest a



restaurant or food shop to open in an area. All we know for now is that there seems to be a preferential difference between what Upper class households prefer and everyone else.

One interesting point to mention is that there are a few outliers which means there may be business opportunity or available market share for the client to gain with these outliers. These outliers can be seen as opportunity as they have been placed into the 'incorrect' cluster due to the lack of certain food businesses that exist within the neighborhood that may push it back into its correct cluster if they were there. This is not concrete but one of the considerations that can come from clustering.

To further develop the analysis, I will be splitting the data in terms of income namely, Upper class, Middle class and Lower class. The three splits were defined with a Median Household Income (MHI) of more than 100 000; between 100 000 and 50 000; and less than 50 000 (all MHI values are in dollars). I will then be building a frequency table using only the 10 most common venues for each Income segment and observe those results.

## 5.2 Frequency Table Results

The following sections show the results from the frequency analysis of food businesses in each MHI bracket.

### 5.2.1 Upper Class areas in New York

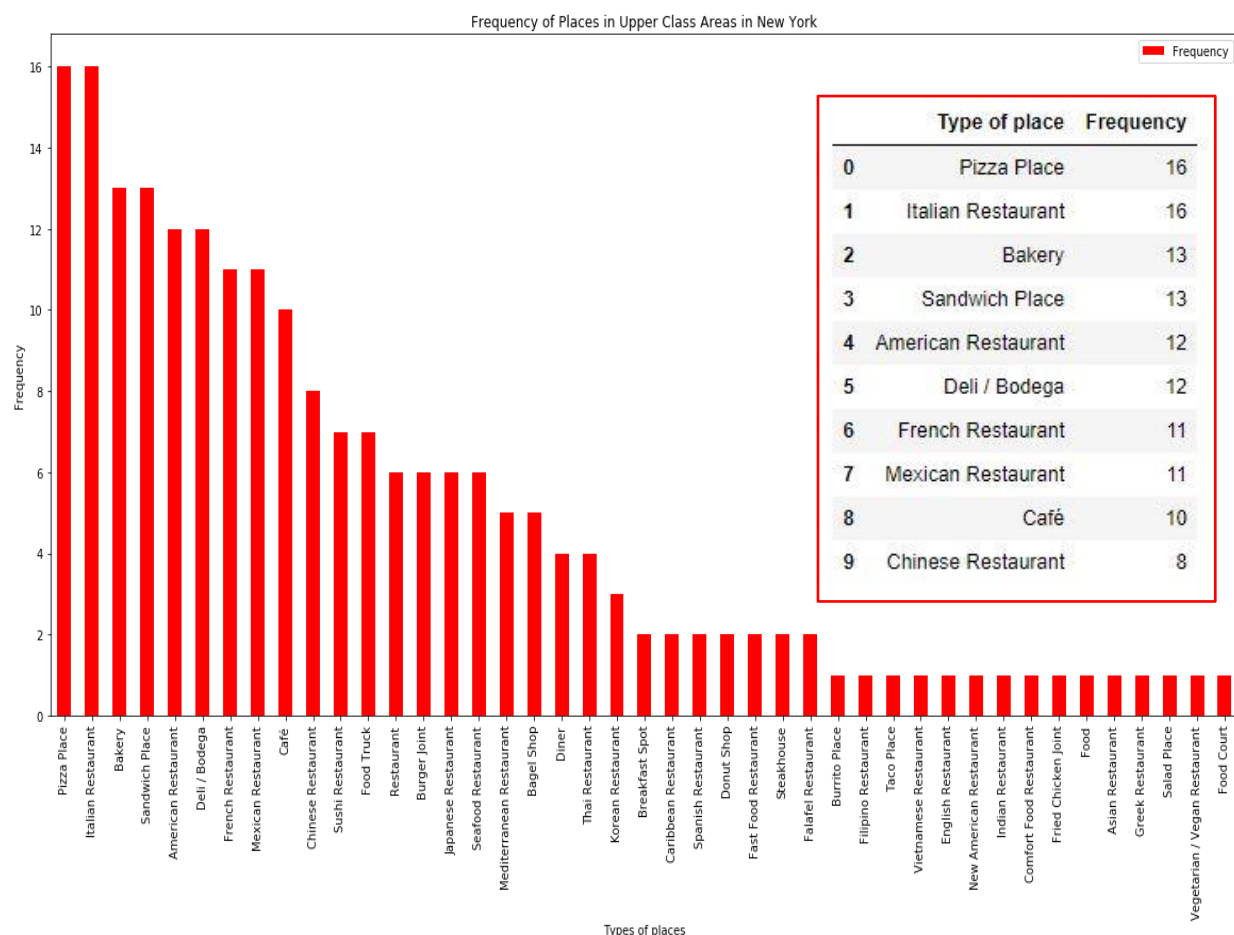


Figure 2: Frequency of Food Businesses in Upper Class areas of New York.

## 5.2.2 Middle Class areas in New York

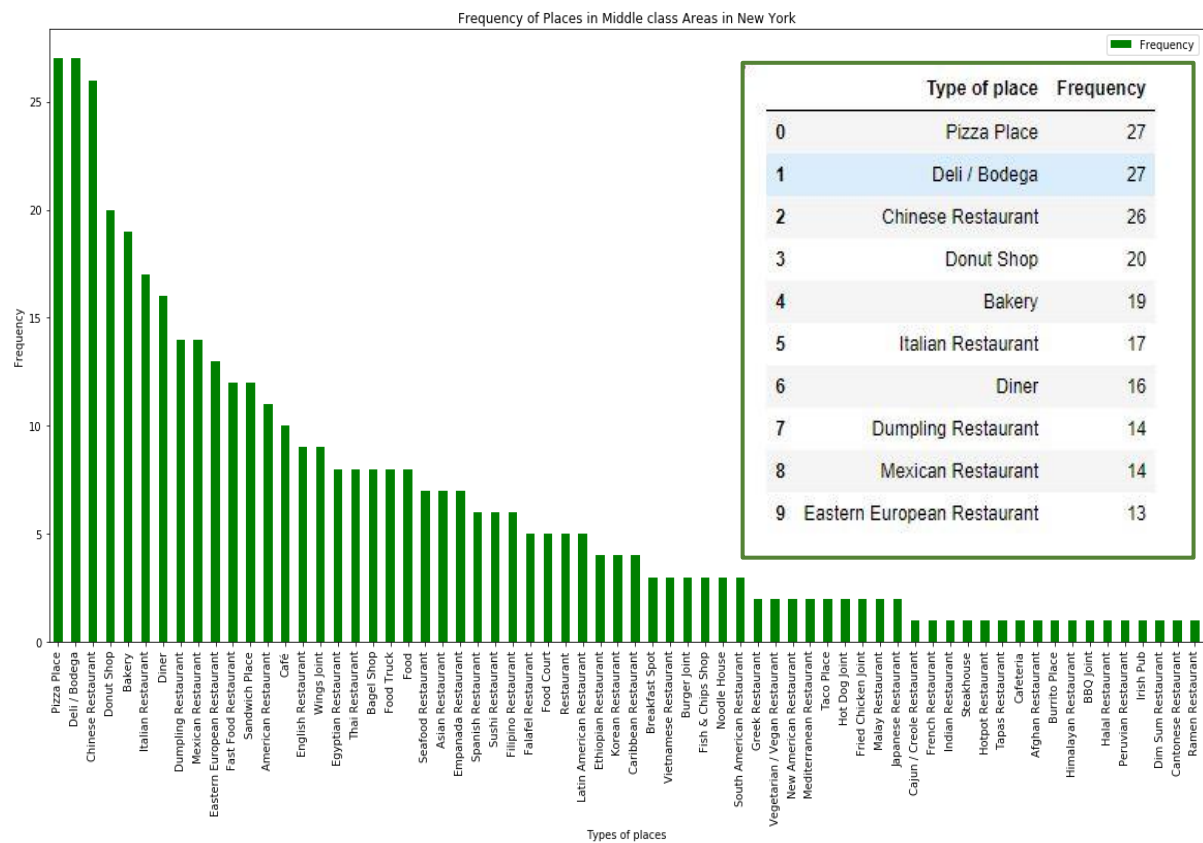


Figure 3: Frequency of Food Businesses in Middle Class areas of New York.

## 5.2.3 Lower Class areas in New York

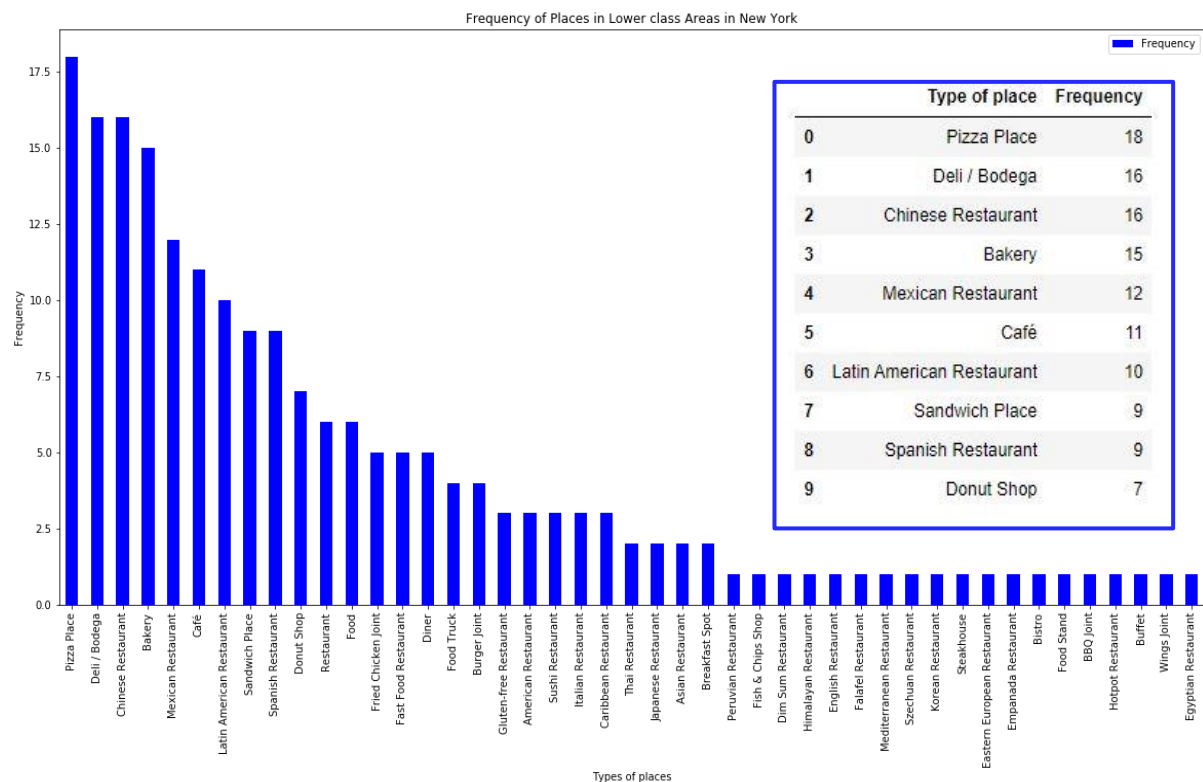


Figure 4:: Frequency of Food Businesses in Lower Class areas of New York.

## 6 Discussion

The following section will discuss each observation that was noticed from the data analysis. This is followed by comments and recommendations.

### Initial Remarks

When making observations with the used data, there are many assumptions that are made. One of the biggest assumptions is that the Demand in the various areas is met by the supply from the various food businesses. It is very possible that these businesses exist and are struggling to keep up with monthly expenses which may be as a result of its economic climate it finds itself in (something that is required to be understood further in the business case development for the client). For this immediate analysis, we assume that these businesses are indeed profitable and that they all are part of fulfilling the demand for types of food venues in their respective areas.

### 6.1 Observation #1

Most importantly to note - *EVERYONE LOVES PIZZA*. Regardless of your income bracket, you will always want and eat pizza. It appears that Pizza businesses cater for all MHI brackets but what would be interesting to see is the quality and price points of these pizza places in their various MHI points.

### 6.2 Observation #2

One of the interesting observations to make is that for the Upper class, Deli/Bodega's does not appear to have as high a demand as it only makes a combined 5th and 6th place in popularity. Living Standards Measure (LSM) is a marketing research tool that groups people according to their living standards and MHI is one of the indexes that help define the LSM. With this in mind, it also suggests that Upper class residents having a living standard that allows them to eat out more often as this is a lifestyle they can easily afford.

### 6.3 Observation #3

Another Observation is that the Upper class does not contain many fast food restaurants as it comes in at a combined 22nd place with only two areas with fast food restaurants in their top 10 most common places. This is surprising as fast food restaurants were expected to be very common across all three MHI splits. It did not come in the top 10 for any MHI bracket with the highest being 11th place for Middle class and combined 13th place for Lower class. It does however combine 13th and 14th place with 'Fried chicken joint' which could be a more specific form of fast food (i.e KFC) which would create a combined frequency of 10 which may push it higher into combined 6th place. Again, deeper analysis to understand this would be to actually look at the various restaurants that were highlighted as 'Fried Chicken Joint' and see if it can be interpreted as Fast food.

### 6.4 Observation #4

Upper neighborhoods have an increased number of sit-down restaurants as oppose to pop in donut shops, fast food restaurants and delis. Which again suggests that if a client wishes to build a sit-down restaurant that contains a specific type of cuisine, the Upper class is most likely the market to move forward with. Otherwise, quick pop in stores that do not necessarily have large seating areas are much more desirable between the Middle and Lower class brackets. Depending on the lavishness of the product (i.e Cupcakes are more lavish than pies), this would determine in which of the final two class areas to move into.

### 6.5 Observation #5

All in all, all food types are mostly accommodated for in every area. The differences that are mentioned provide insight into differences that can be noted to ensure a successful start-up business but in no way has the data shown

that different food types are only found in certain areas. The data in fact suggests that every food business type can be found at any price bracket which somewhat gives freedom for the client to decide which bracket he/she wishes start in and then follow the laid out rules to increase the probability of success.

### **Further Comments**

Another point to make for this discussion is that this research is not absolute and I would recommend further investigation as this data does not allow for sufficient insight to suggest a business venture to someone wishing to open a food business in New York. The data does tell a story but it needs to be verified and investigated further to have a deeper understanding of the New York food business horizon. Many assumptions are made throughout this discussion which is good for initial overview and understanding of this situation but digging deeper makes this insight more concrete or even provide a different picture to what the data is saying.

## **6.6 Recommendations**

- Assumptions need to be verified.
- The quality of restaurants and their price points should be accounted for and trends should be observed.
- Going through the data and seeing if there are overlaps between various categories and observe better ways to categorise businesses.
- Another way to analyse the data would be to have one big request of all the food businesses in New York instead of looking at the radius from the centre of the neighborhood. This would reduce the possibility of overlaps as well as accommodating possible businesses that may have been left out.
- To understand the culture of New Yorkers better, a New Yorker should be brought in that may verify claims and provide a deeper insight as to what he/she thinks about the data or why things may be the way they are.
- what would also be interesting to see is whether restaurants are chain stores in the various areas or independently owned. Depending on the clients wishes, a chain store may be preferable as oppose to starting from scratch but in a business environment where clients prefer the unique experiences of non-chain restaurants, this may be preferable.

Lastly, these recommendations would exist under a perfect scenario without infinite resources. Unfortunately in real life, resources are limited which means that retrieving the above data and analysing it may be an intensive endeavour to embark on in terms of time and money to acquire such data and information. Therefore, it is up to the project manager to liaise with the client to understand the required depth of analysis that is of sufficient standard to make a judgement without depleting resources.

## **7 Conclusion**

Our client has asked our company to investigate the food business environment within New York. This report is a subsection of the investigation looking purely at the businesses that currently exist as well as the MHI of the residents closest to these businesses. The analysis has noted that there are business opportunities across all MHI brackets and that each MHI bracket has a different set of requirements that would increase the likelihood of success in a specific area if a business is opened. Finally, this report is not complete and should be considered in unison with other research that is also conducted in order to better understand the business horizon and hence make a strong business case for the client.