1. INTRODUCTION

Deciding where to place its headquarters is a significant and important decision for any company no matter its size, of course, the bigger the company, the larger, complex and lengthy the problem become.

For this project we will assume that we are part of the Data Science team from a very big overseas Corporation that has decided to open its US Headquarters in New York City. The Board of Directors of this Company now wants to define the precise location for them.

Needless to say, there is a lot to review in this process: legal, logistical, environmental aspects; budgets; timelines, etc. so several departments within the company will be involved. However, all of those will be out of the scope of this project, instead we will focus on getting the insights of New York City neighborhoods that better match the criteria defined by the Board in order to serve optimally the company operations and get the best possible living standards and quality of life for their future employees.

This project will help the Data Science team provide to the Board with a list of neighborhoods that better fit basic criteria, regarding living standards for employees and everyday headquarters operations.

Notes:

-The target audience for this project is the Board of Directors of the Corporation, since we must provide a report to them with the findings of our analysis.

-The goal of this project will be to come up with a list of neighborhoods, not a final one. Further and specific analysis must be done to get one, and that is not possible with a free-limited Foursquare account (Can't get over 100 venues & can't get details of the venues).

2. DATA

For this project we will need to explore, segment and cluster the 5 Boroughs and 306 Neighborhoods of New York City. There is a free dataset on the web that contains the geospatial location of each borough and neighborhoods, i.e. latitude and longitude coordinates of their centroids. The link to this dataset is: https://geo.nyu.edu/catalog/nyu_2451_34572

The format of the dataset is like the following example:

Borough	Neighborhood	Latitude	Longitude
Bronx	Wakefield	40.894705	-73.847201
Bronx	Co-op City	40.874294	-73.829939
Bronx	Eastchester	40.887556	-73.827806

We'll use these coordinates to extract venue information from the Foursquare API, using a free-personal account. We'll find out what are the most common venues in a selected, filtered list of neighborhoods and cluster them according to their similarities (we'll call them TOP neighborhoods).

The dataframe for the most common venues of a neighborhood is going to look like this:

Neighborhood	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3d…
Astoria	1	Middle Eastern Restaurant Bar	Hookah Bar	…
Jackson Heights	1	Latin American Restaurant	Peruvian Restaurant	…

3. METHODOLOGY

A Top-down analysis will be conducted in order to select the best neighborhoods that comply with the minimum requirements defined by the Board of the Company.

First, we'll explore the venues of each of the 5 Boroughs of New York City to find out if they have enough size to locate the Headquarters. In this part, we'll use the amount of neighborhoods that have 50 venues or more as a metric for size: any Borough must have at least 5 neighborhoods with 50+ venues to be considered for the next phase (from now on, we'll refer to this kind of neighborhoods as 'Top Neighborhoods').

Next, we'll explore all the Top Neighborhoods of the selected Boroughs, segmenting them by their 10 most common venues.

Finally, we'll cluster the segmented Top Neighborhoods all across The City and select the one (or ones) that satisfy(ies) our minimum requirements to locate the Headquarters, which are the following:

-Must be commercial area, as neighbors in pure residential areas can be negatively affected.

-High density of hotels nearby for the overnight staying of overseas employees, vendors, visitors, corporate meetings, etc.

-Good number of coffee shops, restaurants and entertainment venues for employees quality of life.

Workflow:

-Using the Geospatial Dataset, different calls requests will be made to the Foursquare API servers to extract information about the venues with the Search feature. Due to the limitations of the free Foursquare account, http requests will be limited on 100 venues and the radius from the neighborhood centroid will be 500 meters.

-Folium/Python visualization library will be used to visualize the neighborhoods cluster distribution of New York City over an interactive leaflet map.

-Unsupervised Machine Learning algorithm K-Means will be applied to form the clusters of different kind of venues from the selected Top Neighborhoods, to reach conclusions.
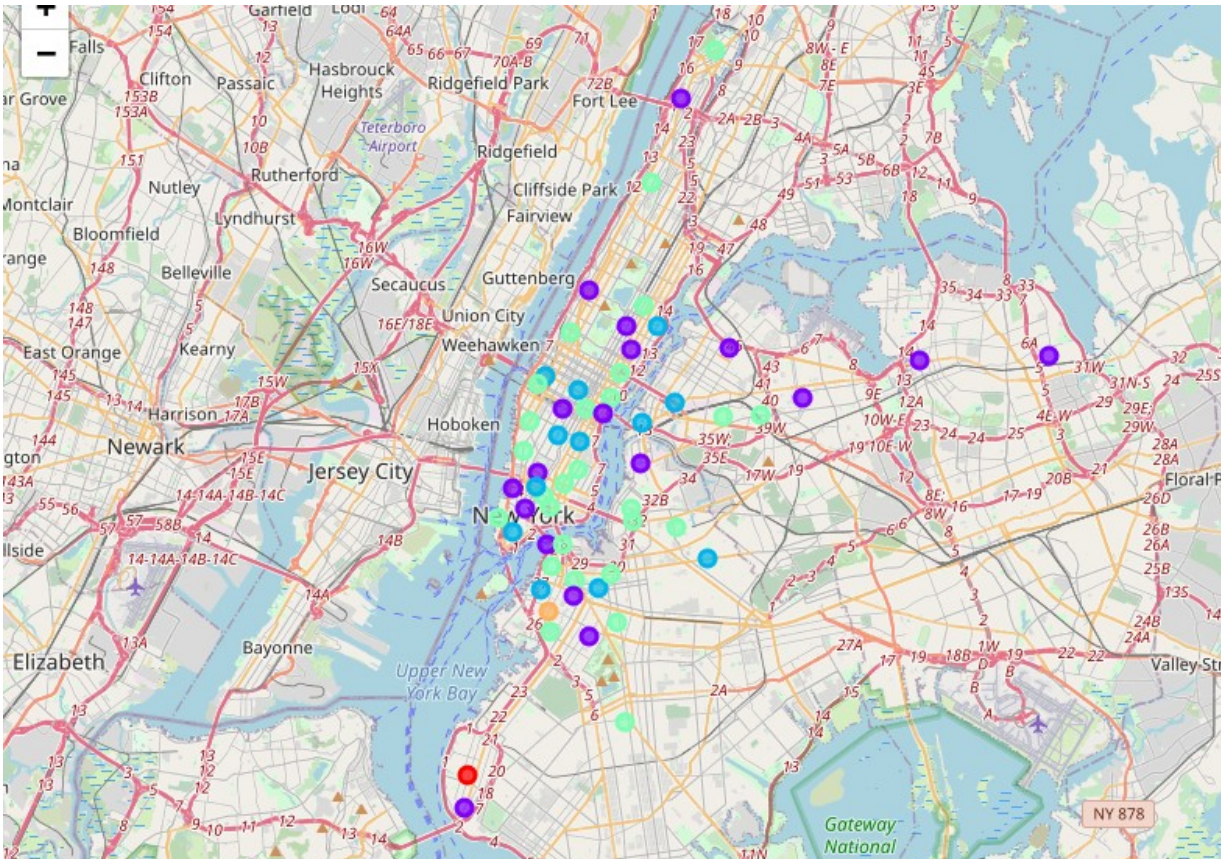
4. RESULTS

In the first part of this Top-down exploratory analysis we got to the following results regarding neighborhoods/venues on each Borough:

Borough	Total amount of Neighborhoods	Total amount of Top Neighborhoods*	% of Top Neighborhoods
Queens	81	9	11.1
Brooklyn	70	20	28.6
Manhattan	40	31	77.5
Bronx	50	3	6.0
Staten Island	63	0	0.0
Total	306	63	20.5

Neighborhoods with 50+ venues.

!! The Boroughs of Bronx and Staten Island do not have the minimum of 5 Top Neighborhoods, so they can not be considered for the next step. **!!**

On the second part of the process, we segmented and clustered the 60 Top Neighborhoods selected by 10 common venues. We got 5 clusters as shown in the next map:

The 5 clusters are shown in the next 5 tables with only the 6 first Most Common Venues, for simplicity:

	Neighborhood	Cluster Label	1st MCV	2nd MCV	3rd MCV	4th MCV	5th MCV	6th MCV
				CLUSTER 0				
0	Bay Ridge	0	Spa	Pizza Place	Italian Restaurant	Greek Restaurant	American Restaurant	Bar

	Neighborhood	Cluster Label	1st MCV	2nd MCV	3rd MCV	4th MCV	5th MCV	6th MCV
				CLUSTER 1				
0	Astoria	1	Middle Eastern Restaurant	Bar	Hookah Bar	Greek Restaurant	Mediterranean Restaurant	Indian Restaurant
1	Jackson Heights	1	Latin American Restaurant	Peruvian Restaurant	South American Restaurant	Bakery	Thai Restaurant	Mobile Phone Shop
2	Flushing	1	Chinese Restaurant	Hotpot Restaurant	Bubble Tea Shop	Karaoke Bar	Korean Restaurant	Bakery
3	Bayside	1	Bar	Indian Restaurant	Pizza Place	Sushi Restaurant	American Restaurant	Pub
4	Washington Heights	1	Café	Bakery	Grocery Store	Chinese Restaurant	Mobile Phone Shop	Pizza Place
5	Upper East Side	1	Italian Restaurant	Juice Bar	Gym / Fitness Center	Exhibit	Bakery	Yoga Studio
6	Lenox Hill	1	Italian Restaurant	Pizza Place	Coffee Shop	Sushi Restaurant	Cocktail Bar	Café
7	Upper West Side	1	Italian Restaurant	Bakery	Wine Bar	Coffee Shop	Dessert Shop	Bagel Shop
8	Greenwich Village	1	Italian Restaurant	Coffee Shop	Café	Bakery	Wine Bar	Gym
9	Tribeca	1	Park	Italian Restaurant	Café	Wine Bar	Spa	Bakery
10	Civic Center	1	Hotel	Coffee Shop	Cocktail Bar	Spa	Yoga Studio	French Restaurant
11	Midtown South	1	Korean Restaurant	Hotel	Japanese Restaurant	Burger Joint	Café	Dessert Shop
12	Tudor City	1	Café	Park	Mexican Restaurant	Deli / Bodega	Pizza Place	Thai Restaurant
13	Greenpoint	1	Bar	Pizza Place	Cocktail Bar	Coffee Shop	Café	Sushi Restaurant
14	Park Slope	1	Coffee Shop	Burger Joint	American Restaurant	Pub	Pet Store	Bagel Shop
15	Boerum Hill	1	Coffee Shop	Bar	Dance Studio	French Restaurant	Spa	Arts & Crafts Store
16	Fort Hamilton	1	Sandwich Place	Deli / Bodega	Bank	Pharmacy	Chinese Restaurant	Bagel Shop
17	Fulton Ferry	1	Park	American Restaurant	Scenic Lookout	Bakery	Ice Cream Shop	Roof Deck

Cluster 2

	Neighborhood	Cluster Label	1st MCV	2nd MCV	3rd MCV	4th MCV	5th MCV	6th MCV
0	Long Island City	2	Hotel	Coffee Shop	Pizza Place	Bar	Mexican Restaurant	Café
1	Rockaway Beach	2	Beach	Ice Cream Shop	BBQ Joint	Pizza Place	Seafood Restaurant	Food Truck
2	Hunters Point	2	Café	Japanese Restaurant	Italian Restaurant	Wine Shop	Scenic Lookout	American Restaurant
3	Yorkville	2	Coffee Shop	Italian Restaurant	Gym	Bar	Deli / Bodega	Sushi Restaurant
4	Clinton	2	Theater	Gym / Fitness Center	Coffee Shop	Wine Shop	Gym	Hotel
5	Midtown	2	Hotel	Clothing Store	Coffee Shop	Theater	Cuban Restaurant	Pizza Place
6	Soho	2	Italian Restaurant	Coffee Shop	Mediterranean Restaurant	French Restaurant	Spa	Gym
7	Gramercy	2	Coffee Shop	Italian Restaurant	Grocery Store	Pizza Place	Playground	Bar
8	Financial District	2	Coffee Shop	Hotel	Gym / Fitness Center	Salad Place	Cocktail Bar	Park
9	Flatiron	2	Gym / Fitness Center	Yoga Studio	Outdoor Sculpture	Cosmetics Shop	Park	Wine Shop
10	Bushwick	2	Bar	Mexican Restaurant	Deli / Bodega	Coffee Shop	Pizza Place	Bakery
11	Cobble Hill	2	Playground	Coffee Shop	Bar	Pizza Place	Yoga Studio	Cocktail Bar
12	Fort Greene	2	Italian Restaurant	Flower Shop	Wine Shop	Coffee Shop	Yoga Studio	Theater

Cluster 3

	Neighborhood	Cluster Label	1st MCV	2nd MCV	3rd MCV	4th MCV	5th MCV	6th MCV
0	Woodside	3	Grocery Store	Bakery	Thai Restaurant	Latin American Restaurant	Filipino Restaurant	American Restaurant
1	Sunnyside Gardens	3	Bar	Grocery Store	Pizza Place	Thai Restaurant	Coffee Shop	Pharmacy
2	Chinatown	3	Chinese Restaurant	Cocktail Bar	Bakery	Salon / Barbershop	Coffee Shop	Optical Shop
3	Inwood	3	Mexican Restaurant	Café	Bakery	Lounge	Pizza Place	Restaurant
4	Hamilton Heights	3	Pizza Place	Café	Coffee Shop	Deli / Bodega	Mexican Restaurant	Yoga Studio
5	Lincoln Square	3	Gym / Fitness Center	Plaza	Café	Theater	Italian Restaurant	Concert Hall
6	Murray Hill	3	Sandwich Place	Japanese Restaurant	Coffee Shop	Burger Joint	Steakhouse	Pizza Place
7	Chelsea	3	Art Gallery	Coffee Shop	Seafood Restaurant	Ice Cream Shop	Italian Restaurant	Park
8	East Village	3	Pizza Place	Coffee Shop	Juice Bar	Japanese Restaurant	Mexican Restaurant	Vietnamese Restaurant
9	Little Italy	3	Chinese Restaurant	Italian Restaurant	Spa	Pizza Place	Bubble Tea Shop	Bakery
10	West Village	3	Wine Bar	Italian Restaurant	Park	Coffee Shop	American Restaurant	New American Restaurant
11	Battery Park City	3	Park	Hotel	Gym	Memorial Site	Boat or Ferry	Wine Shop
12	Carnegie Hill	3	Coffee Shop	Café	Yoga Studio	Bar	Italian Restaurant	Gym / Fitness Center
13	Noho	3	Italian Restaurant	Coffee Shop	Grocery Store	Seafood Restaurant	Japanese Restaurant	Bar
14	Sutton Place	3	Italian Restaurant	Gym / Fitness Center	Park	Coffee Shop	Bagel Shop	Yoga Studio
15	Turtle Bay	3	Italian Restaurant	Coffee Shop	Park	Deli / Bodega	Wine Bar	French Restaurant
16	Hudson Yards	3	Hotel	American Restaurant	Gym / Fitness Center	Italian Restaurant	Dog Run	Gym
17	Prospect Heights	3	Bar	Mexican Restaurant	Cocktail Bar	Ice Cream Shop	Coffee Shop	Beer Bar
18	Brooklyn Heights	3	Yoga Studio	Deli / Bodega	Park	Bakery	Gym	Cosmetics Shop
19	Gowanus	3	Art Gallery	Italian Restaurant	Furniture / Home Store	Bar	Deli / Bodega	Wine Shop
20	Clinton Hill	3	Pizza Place	Italian Restaurant	Wine Shop	Mexican Restaurant	Thai Restaurant	Yoga Studio
21	Downtown	3	Burger Joint	Sandwich Place	Pizza Place	Coffee Shop	Chinese Restaurant	Bakery
22	East Williamsburg	3	Deli / Bodega	Bar	Cocktail Bar	Bakery	Coffee Shop	Concert Hall
23	North Side	3	Coffee Shop	Pizza Place	Yoga Studio	American Restaurant	Bar	Wine Bar
24	South Side	3	Bar	Pizza Place	Coffee Shop	American Restaurant	Yoga Studio	Breakfast Spot
25	Ditmas Park	3	Donut Shop	Caribbean Restaurant	Chinese Restaurant	Pizza Place	Discount Store	Yoga Studio
26	Dumbo	3	Park	Coffee Shop	Yoga Studio	Boxing Gym	Café	Scenic Lookout

Cluster 4

	Neighborhood	Cluster Label	1st MCV	2nd MCV	3rd MCV	4th MCV	5th MCV	6th MCV
0	Carroll Gardens	4	Italian Restaurant	Coffee Shop	Pizza Place	Cocktail Bar	Bakery	Wine Shop

5. DISCUSSION

The clustering process gave us two large, relevant, homogeneous clusters to compare: Cluster 2 and Cluster 3.

We cross examined those clusters with the Minimum Requirements for a Top Neighborhood to be selected as location for Headquarters:

-Must be commercial area, as neighbors in pure residential areas can be negatively affected.

-High density of hotels nearby for the overnight staying of overseas employees, vendors, visitors, corporate meetings, etc.

-Good number of coffee shops, restaurants and entertainment venues for employees quality of life.

We can clearly see that cluster number 2 better fit them.

6. CONCLUSION

!! So the final report to the Board of Directors of the Company, considering a recommended location for US Headquarters in New York City will list the following neighborhoods: **!!**

Neighborhood	Borough
Long Island City	Queens
Rockaway Beach	Queens
Hunters Point	Queens
Yorkville	Manhattan
Clinton	Manhattan
Midtown	Manhattan
Soho	Manhattan
Gramercy	Manhattan
Financial District	Manhattan
Flatiron	Manhattan
Bushwick	Brooklyn
Cobble Hill	Brooklyn
Fort Greene	Brooklyn

*Interestingly enough to mention, on 2019 the online retailer Amazon initially selected the neighborhood of Long Island City, Queens for its HQ2 project, which is the first one in our list!