

The Battle of Neighborhoods

"Entering international palates"

NewYork - Manhattan



INTRODUCTION

The restaurant chain has been in the local market for 20 years and has around 150 franchises distributed throughout the local territory

As part of the internalization strategy, investors want to assess the risks and their chances of success in this new adventure



BUSINESS PROBLEM

Manhattan is an island located at the mouth of the Hudson River in northern New York Harbor and is also one of five districts that are part of New York City, United States

90% of New York's attractions are concentrated in Manhattan.

The consultant does not know the different types of business and the distribution of these in the different neighborhoods of the borough

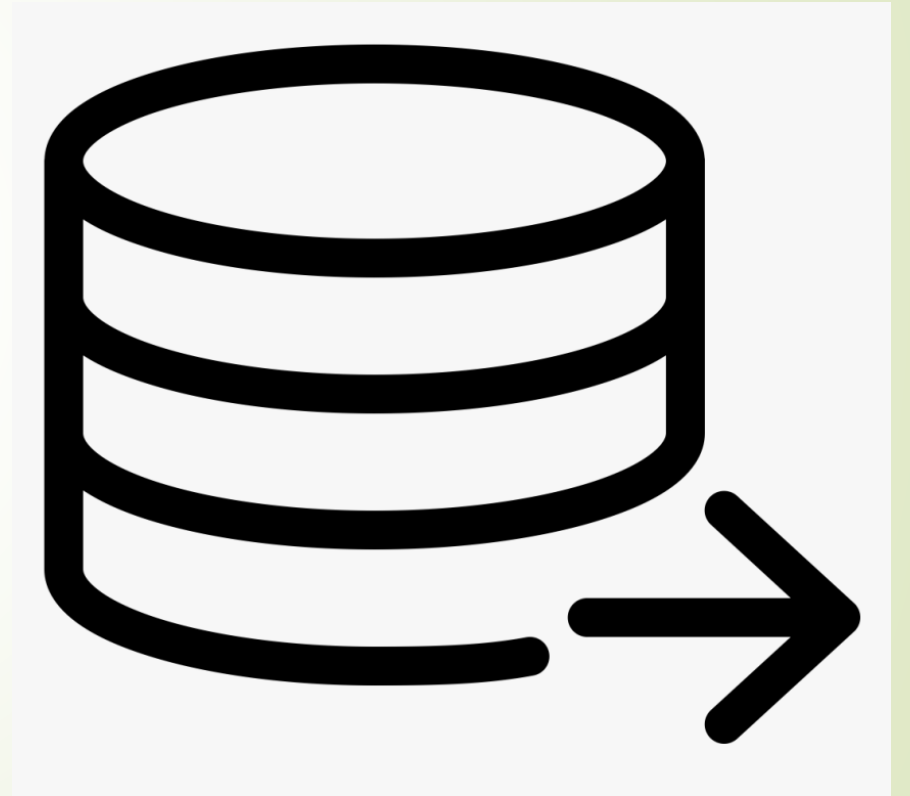
Problem. How are the different types of businesses distributed in the city's neighborhoods to guide market research on the most important ones based on the distribution of places or commercial businesses?



DATA

The commercial businesses of the different Manhattan neighborhoods will be obtained from the Foursquare API: <https://es.foursquare.com/>

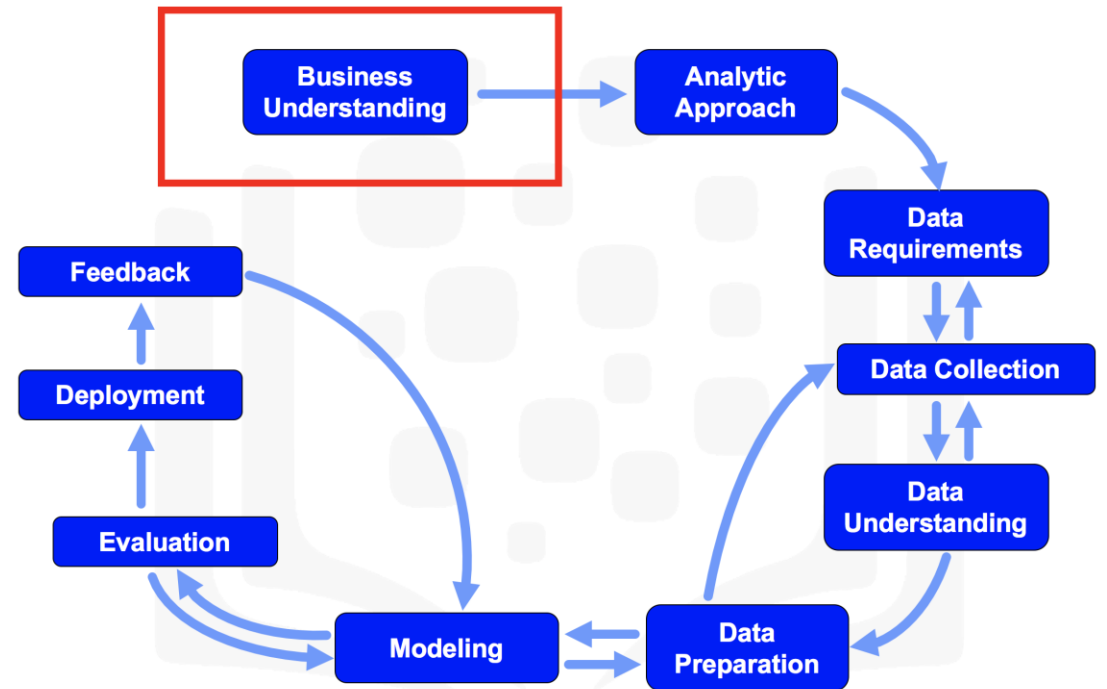
As part of obtaining the data mentioned above, we have to obtain the coordinates of the neighborhoods for them we will use the following source: https://geo.nyu.edu/catalog/nyu_2451_34572



METHODOLOGY

Business Understanding: How are the different types of businesses distributed in the borough's neighborhoods to guide market research on the most important ones?

Analitic Approach: To answer the question, the machine leaning approach was established with the K means algorithm

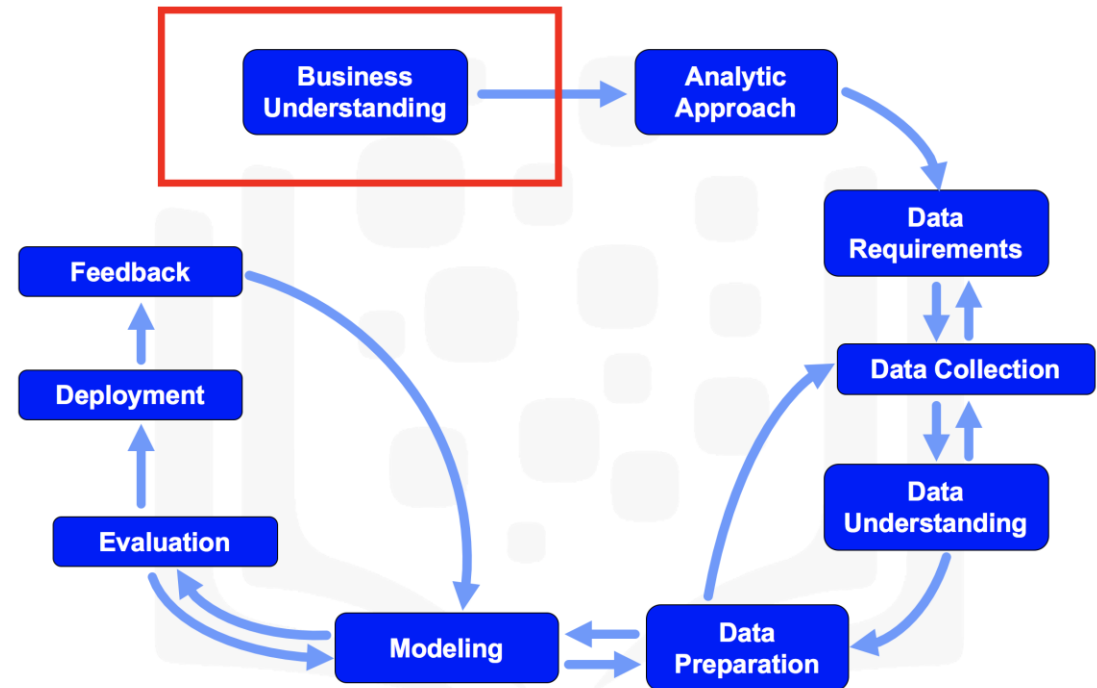


METHODOLOGY

Data Requeriments, Collections, Understanding and Preparation:

- The data requirements were to obtain data from the Manhattan borough mainly: Neighborhoods, categories of venues with coordinates.
- 2 dataset were identified for our case study: a New York geojson file and information about the venues around the neighborhoods from the Foursquare API.
- Finally, the preparation of the data resulted in a dataset that contained each neighborhood with the average of the number of venues around them.

Modeling: For the data modeling, the K means algorithm was established and an initial number of 5 clusters was defined.



RESULTS

Cluster 1 (1 neighborhood)

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
Midtown South	Korean Restaurant	Hotel	Japanese Restaurant	Dessert Shop	Burger Joint	Gym / Fitness Center	Clothing Store	Bar	American Restaurant	Coffee Shop

Cluster 2 (7 neighborhood)

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
Marble Hill	Sandwich Place	Coffee Shop	Gym	Ice Cream Shop	Department Store	Pharmacy	Diner	Discount Store	Donut Shop	Bank
Roosevelt Island	Park	Pizza Place	Sandwich Place	Gym / Fitness Center	Gym	Coffee Shop	Liquor Store	Outdoors & Recreation	Dry Cleaner	School
Midtown	Coffee Shop	Hotel	Sporting Goods Shop	Clothing Store	Steakhouse	Spa	Bakery	Café	Sandwich Place	Theater
Murray Hill	Sandwich Place	Hotel	Coffee Shop	Sushi Restaurant	Bar	Japanese Restaurant	Burger Joint	Gym / Fitness Center	American Restaurant	Cuban Restaurant
Morningside Heights	Park	Bookstore	American Restaurant	Coffee Shop	Deli / Bodega	Burger Joint	Pub	Farmers Market	Mediterranean Restaurant	Supermarket

RESULTS

Cluster 3 (22 neighborhood)

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
Manhattanville	Coffee Shop	Seafood Restaurant	Liquor Store	Italian Restaurant	Deli / Bodega	Park	Mexican Restaurant	Bus Stop	Bus Station	Spanish Restaurant
Upper East Side	Italian Restaurant	Coffee Shop	Bakery	Exhibit	Spa	French Restaurant	Juice Bar	Yoga Studio	Gym / Fitness Center	Wine Shop
Yorkville	Italian Restaurant	Coffee Shop	Gym	Bar	Deli / Bodega	Pizza Place	Sushi Restaurant	Diner	Wine Shop	Mexican Restaurant
Lenox Hill	Italian Restaurant	Coffee Shop	Sushi Restaurant	Pizza Place	Cocktail Bar	Café	Gym	Gym / Fitness Center	Burger Joint	Turkish Restaurant
Upper West Side	Italian Restaurant	Coffee Shop	Bar	Bakery	Pizza Place	Ice Cream Shop	Café	Seafood Restaurant	Shoe Store	Bookstore

RESULTS

Cluster 4 (9 neighborhood)

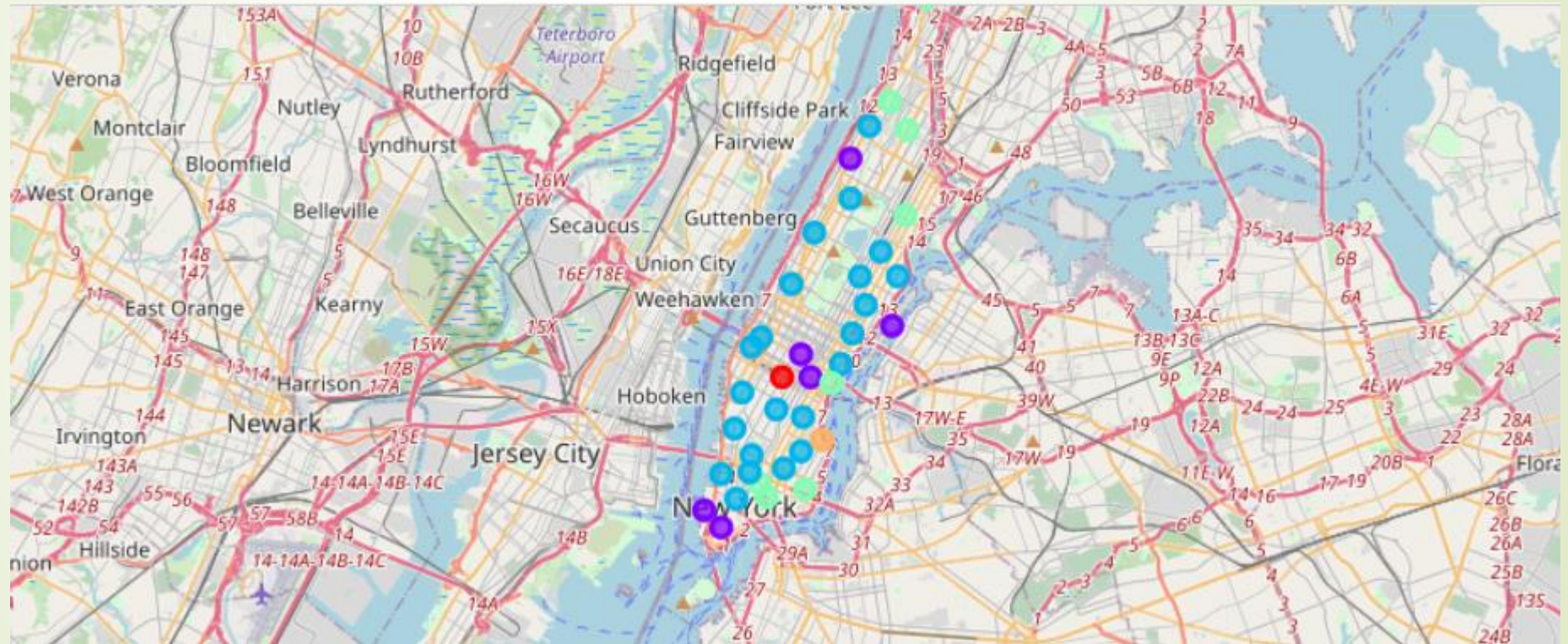
Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
Chinatown	Chinese Restaurant	Cocktail Bar	American Restaurant	Bakery	Spa	Dessert Shop	Optical Shop	Hotpot Restaurant	Vietnamese Restaurant	Salon / Barbershop
Washington Heights	Café	Bakery	Grocery Store	Chinese Restaurant	Mobile Phone Shop	Deli / Bodega	Coffee Shop	Supermarket	Supplement Shop	Tapas Restaurant
Inwood	Mexican Restaurant	Pizza Place	Restaurant	Café	Lounge	Frozen Yogurt Shop	Bakery	Park	Chinese Restaurant	Deli / Bodega
Hamilton Heights	Pizza Place	Café	Coffee Shop	Deli / Bodega	Mexican Restaurant	Bakery	Sushi Restaurant	Cocktail Bar	Sandwich Place	Yoga Studio
Central Harlem	Gym / Fitness Center	French Restaurant	Fried Chicken Joint	African Restaurant	American Restaurant	Bar	Seafood Restaurant	Chinese Restaurant	Art Gallery	Bookstore

Cluster 5 (1 neighborhood)

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
Stuyvesant Town	Boat or Ferry	Park	Bar	Pet Service	German Restaurant	Baseball Field	Fountain	Harbor / Marina	Cocktail Bar	Coffee Shop

RESULTS

Cluster distribution)





DISCUSSION

In relation to the results presented by the 5 clusters associated with its particular development of shops and places at the end, we observed Clusters 3 as the most representative of the Manhattan district. Just looking at Cluster 3, there are 80% of restaurants, manifesting the large number of competitors in the Cluster neighborhoods.

On the other hand, observing Clusters 2 and 4 as we observe that the percentage is less than 40% but it is complemented by fast food places and coffee.



CONCLUSION

Getting to know which neighborhoods are more attractive based on the different commercial and cultural places, among others, provides us with a perspective focused on the services that are developed in the neighborhood and being a starting point to guide marketing efforts in the environment.

In addition, according to the results of the data modeling, it has oriented us to direct the efforts of searching and deepening the analysis in those neighborhoods with similarities and that may have a greater social-economic impact in the Manhattan district (Cluster 3).

Finally, this first approximation of neighborhoods does not give the opportunity to make a segmentation of neighborhoods based on other aspects, such as food places and associated market shares.