

Capstone Project – The Battle of Neighborhoods

By
Manuel Martinez



DESCRIPTION OF THE PROBLEM

- The Manhattan district is one of the best known and most visited in New York, in which more than 90% of the main actions of the city are concentrated: the Empire State Building, the Rockefeller Center, the Chrysler Building and the incredible Times Square, to name a few.
- In addition to its wonderful tourist and commercial attraction, it is an ideal place to live, surrounded by the Hudson (west) and Harlem (north) rivers. Each person is different, with different tastes and interests, and these are really important factors when buying an apartment. It is also necessary to know the environment and the places that are in the surroundings of the new home, since these provide an overview of the quality of life that can be had there and is a very valuable element to promote the purchase of real estate goods.

That is why a person before buying an apartment in Manhattan asks the following question:

**how to consider and choose the best neighborhood to live in
Manhattan according to the main places that surround the area?**

DATA



Two different sources were taken as reference for the elaboration of this project:

- The first comes from the renowned Kaggle site. A record was selected for each building or construction unit (apartment, residence, etc.) sold on the New York City property market, over a 12-month period, specifically September 2016 to the same month in 2017.

<https://www.kaggle.com/new-york-city/nyc-property-sales>

- The second source was taken from Foursquare, through its developer API. Foursquare is a specialized location technology platform dedicated to improving the way people move around the real world. It is a service based on web location applied to social networks. From it we drew the busiest places around the Manhattan district neighborhoods.

<https://es.foursquare.com/>

METHODOLOGY

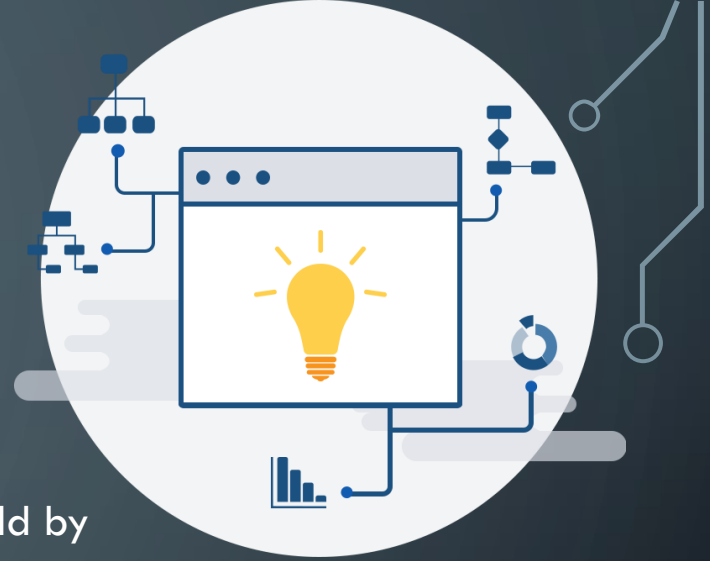
- **Data processing and Model Creation**

Once the data was loaded, it was advanced in its processing in order to obtain the information without elements that could affect it. Fortunately, the selected data set was already processed for the use we would give it.

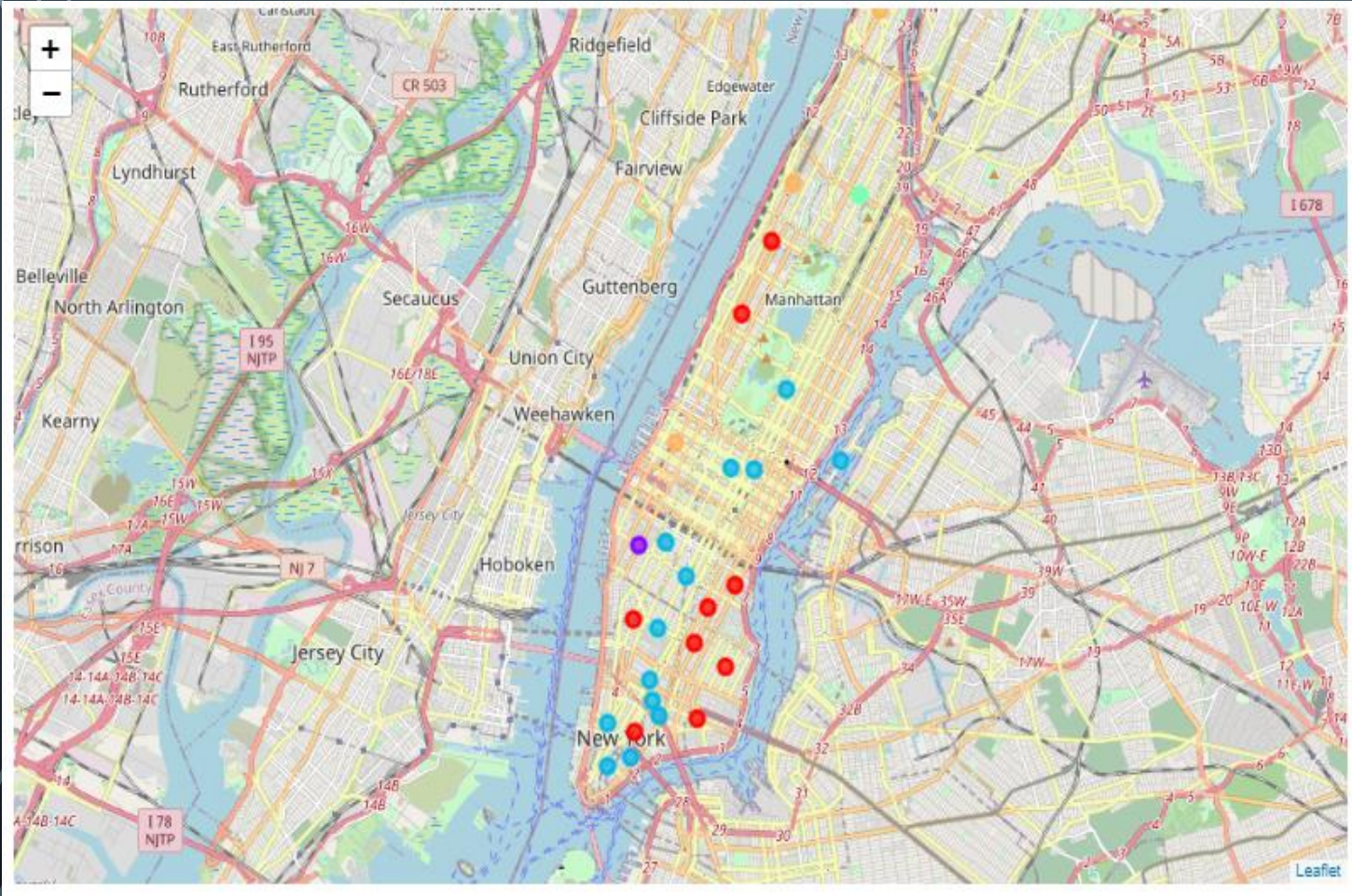
Unnamed: 0	BOROUGH	NEIGHBORHOOD	BUILDING CLASS CATEGORY	TAX CLASS AT PRESENT	BLOCK	LOT	EASE-MENT	BUILDING CLASS AT PRESENT	ADDRESS	APARTMENT NUMBER	ZIP CODE	RESIDENTIAL UNITS	COM
0	4	1.0	ALPHABET CITY	07 RENTALS - WALKUP APARTMENTS	2A	392.0	6.0	C2	153 AVENUE B		10009.0	5.0	
1	5	1.0	ALPHABET CITY	07 RENTALS - WALKUP APARTMENTS	2	399.0	26.0	C7	234 EAST 4TH STREET		10009.0	28.0	
2	6	1.0	ALPHABET CITY	07 RENTALS - WALKUP APARTMENTS	2	399.0	39.0	C7	197 EAST 3RD STREET		10009.0	16.0	
3	7	1.0	ALPHABET CITY	07 RENTALS - WALKUP APARTMENTS	2B	402.0	21.0	C4	154 EAST 7TH STREET		10009.0	10.0	
4	8	1.0	ALPHABET CITY	07 RENTALS - WALKUP APARTMENTS	2A	404.0	55.0	C2	301 EAST 10TH STREET		10009.0	6.0	

Steps

1. Borough and neighborhoods columns were selected
2. Filter to work with the manhattan neighborhood (number 1)
3. It was grouped by neighborhood and a column was created with the houses sold by neighborhood
4. The names of some neighborhoods were arranged and a dataframe was created with the coordinates of each of the neighborhoods
5. The two dataframes were merged
6. Some neighborhoods not belonging to the studied borough were eliminated
7. Some neighborhoods that did not belong to the studied borough were eliminated (Data errors)
8. Connection was made with the foursquare api to get the main venues of each neighborhood
9. Categorical variables were converted to numerical variables using dummy variables and grouped by neighborhood
10. The KMeans model was created to cluster the neighborhoods according to the venues categories



RESULTS AND DISCUSSIONS



After completing the modeling, which resulted in five clusters each with the top 10 venues locations in each neighborhood, they were categorized as follows:

RESULTS AND DISCUSSIONS

	NEIGHBORHOOD	HOUSES SOLD	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	ALPHABET CITY	139	Cocktail Bar	Bar	Wine Bar	Coffee Shop	Garden	Italian Restaurant	Salon / Barbershop	Pizza Place	Beer Bar	Bookstore
3	CIVIC CENTER	296	Chinese Restaurant	Bubble Tea Shop	Dim Sum Restaurant	Park	Coffee Shop	Dessert Shop	Gym	Optical Shop	Cocktail Bar	Cantonese Restaurant
4	EAST VILLAGE	105	Japanese Restaurant	Grocery Store	Dessert Shop	Pizza Place	Bar	Vietnamese Restaurant	Sushi Restaurant	Vegetarian / Vegan Restaurant	Ice Cream Shop	Coffee Shop
8	GRAMERCY	167	Italian Restaurant	American Restaurant	Pizza Place	Bar	Wine Shop	Bagel Shop	Coffee Shop	Spa	Grocery Store	Hotel
9	GREENWICH VILLAGE-CENTRAL	247	American Restaurant	Italian Restaurant	Coffee Shop	Cocktail Bar	Pizza Place	Bakery	Sandwich Place	Speakeasy	Jazz Club	Ice Cream Shop
16	KIPS BAY	303	Bar	Ice Cream Shop	American Restaurant	Bagel Shop	Italian Restaurant	Grocery Store	Coffee Shop	Pizza Place	Yoga Studio	Convenience Store
18	LOWER EAST SIDE	126	Mexican Restaurant	Café	Cocktail Bar	Coffee Shop	American Restaurant	Ice Cream Shop	Sandwich Place	Chinese Restaurant	Bakery	Bar
19	MANHATTAN VALLEY	193	Chinese Restaurant	Coffee Shop	Bar	Yoga Studio	Bistro	Bubble Tea Shop	American Restaurant	Grocery Store	Bagel Shop	Bank
29	UPPER WEST SIDE	1350	American Restaurant	Bar	Coffee Shop	Wine Bar	Dessert Shop	Bakery	Italian Restaurant	Pizza Place	Seafood Restaurant	Bagel Shop

- **Cluster 1:** in these neighborhoods the most frequented places correspond to restaurants and bars, which could be an option for singles, young couples or foodies.

- **Cluster 2:** The main places that surround these neighborhoods are conformed by art galleries and cafeterias. It could be an alternative for young university students, singles or adults who enjoy the arts and the wonders of a quiet life.

	NEIGHBORHOOD	HOUSES SOLD	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	CHELSEA	487	Art Gallery	Café	Gym / Fitness Center	Thai Restaurant	Ice Cream Shop	Grocery Store	Park	Coffee Shop	Bagel Shop	American Restaurant

RESULTS AND DISCUSSIONS

	NEIGHBORHOOD	HOUSES SOLD	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	CHINATOWN	114	Chinese Restaurant	Bakery	Bubble Tea Shop	Sandwich Place	Salon / Barbershop	Vietnamese Restaurant	Spa	Ice Cream Shop	Optical Shop	Cocktail
5	FASHION	92	Coffee Shop	Gym / Fitness Center	Hotel	Boxing Gym	Grocery Store	Lounge	Yoga Studio	Bakery	Donut Shop	Miscellaneous
6	FINANCIAL	470	Coffee Shop	Pizza Place	American Restaurant	Gym	Italian Restaurant	Gym / Fitness Center	Steakhouse	Bar	Falafel Restaurant	Juice
7	FLATIRON	345	Japanese Restaurant	Spa	American Restaurant	Gym / Fitness Center	Gym	Café	Italian Restaurant	Mediterranean Restaurant	Cosmetics Shop	Juice
10	GREENWICH VILLAGE-WEST	363	American Restaurant	Coffee Shop	Yoga Studio	Steakhouse	Ice Cream Shop	Gym / Fitness Center	Italian Restaurant	Sandwich Place	Cocktail Bar	Salad Place
17	LITTLE ITALY	37	Spa	Chinese Restaurant	Ice Cream Shop	Bubble Tea Shop	Mediterranean Restaurant	Bakery	Italian Restaurant	Thai Restaurant	Pizza Place	Sandwich Place
20	MIDTOWN	154	Hotel	Indian Restaurant	Food & Drink Shop	Theater	Sushi Restaurant	Hotel Bar	Art Museum	Vegetarian / Vegan Restaurant	Steakhouse	Pizza
21	MIDTOWN EAST	671	Hotel	Coffee Shop	Salon / Barbershop	Boutique	Art Museum	Gym / Fitness Center	Jewelry Store	Sandwich Place	Chinese Restaurant	Spa
24	ROOSEVELT ISLAND	19	Baseball Field	Park	Playground	Scenic Lookout	Restaurant	Dog Run	Liquor Store	Supermarket	Sandwich Place	School
25	SOHO	205	Italian Restaurant	Mediterranean Restaurant	Coffee Shop	Sandwich Place	Ice Cream Shop	Bakery	Pizza Place	Cosmetics Shop	Clothing Store	Salon / Barbershop
26	SOUTHBRIDGE	60	Italian Restaurant	American Restaurant	Café	Coffee Shop	Sandwich Place	Pizza Place	Hotel	Falafel Restaurant	Juice Bar	
27	TRIBECA	617	Gym / Fitness Center	Coffee Shop	Italian Restaurant	Bakery	Gym	Spa	American Restaurant	Steakhouse	Sushi Restaurant	Cocktail
28	UPPER EAST SIDE	1387	Italian Restaurant	Art Museum	Clothing Store	French Restaurant	Outdoor Sculpture	Coffee Shop	Boutique	Women's Store	Café	Steakhouse

- **Cluster 3:** These neighborhoods have very crowded places around them and are mainly characterized by the diversity between them. Ideal for professionals or contemporary couples with high financial income and whose routine is quite dynamic.

RESULTS AND DISCUSSIONS

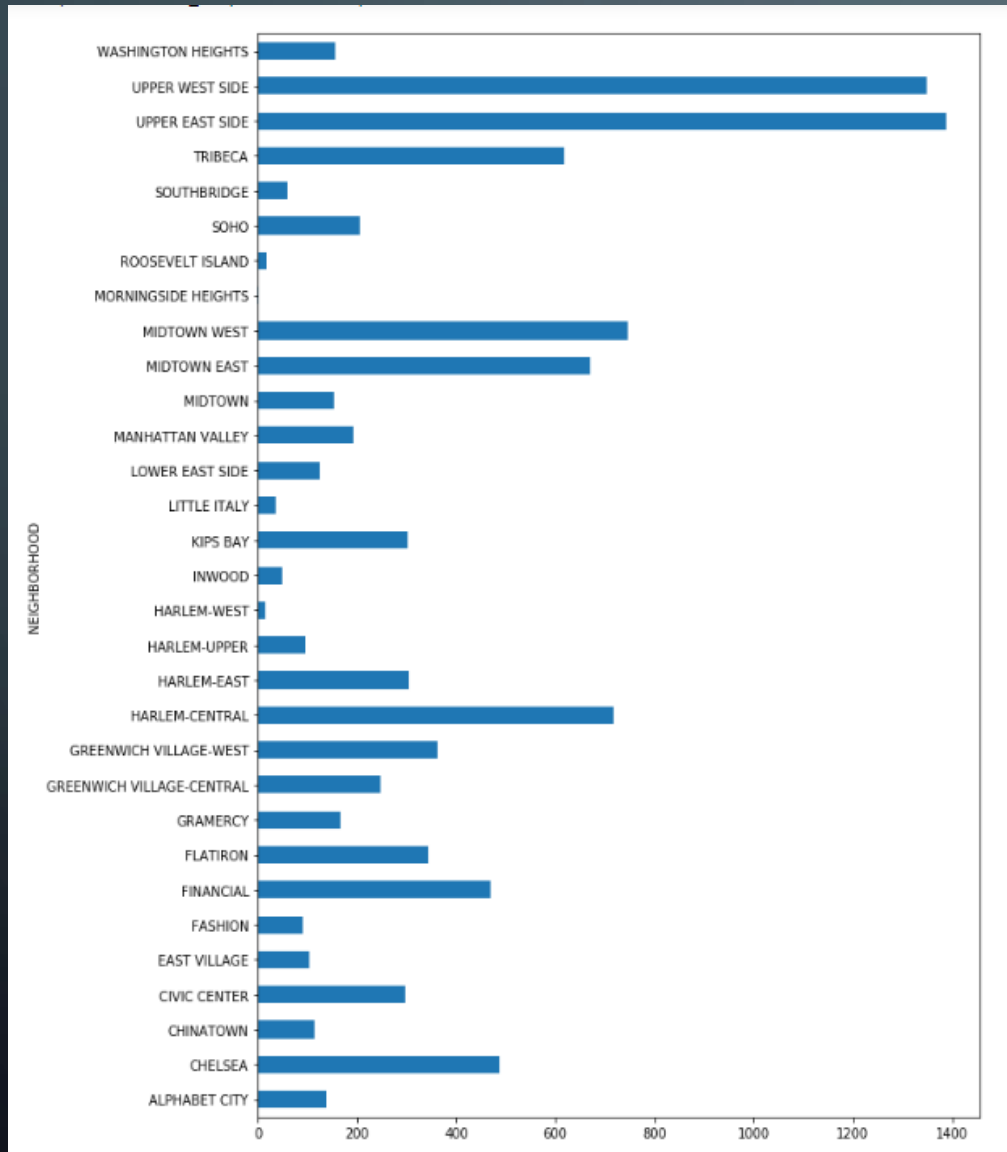
- **Cluster 4:** The busiest places in these neighborhoods are primarily beauty salons, clothing stores, tech shops, and theme restaurants. An ideal option for professional women who pursue a good lifestyle and enjoy international cuisine.

	NEIGHBORHOOD	HOUSES SOLD	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
11	HARLEM-CENTRAL	718	Cosmetics Shop	Mobile Phone Shop	Clothing Store	African Restaurant	Burger Joint	Theater	Southern / Soul Food Restaurant	Pizza Place	Mexican Restaurant	French Restaurant
12	HARLEM-EAST	308	Cosmetics Shop	Mobile Phone Shop	Clothing Store	African Restaurant	Burger Joint	Theater	Southern / Soul Food Restaurant	Pizza Place	Mexican Restaurant	French Restaurant
13	HARLEM-UPPER	97	Cosmetics Shop	Mobile Phone Shop	Clothing Store	African Restaurant	Burger Joint	Theater	Southern / Soul Food Restaurant	Pizza Place	Mexican Restaurant	French Restaurant
14	HARLEM-WEST	14	Cosmetics Shop	Mobile Phone Shop	Clothing Store	African Restaurant	Burger Joint	Theater	Southern / Soul Food Restaurant	Pizza Place	Mexican Restaurant	French Restaurant

- **Cluster 5:** The main places that surround these neighborhoods are made up of parks, cafes, bookstores and wine shops. It could be an alternative for families with children, retired adults who enjoy outdoor walks or a wine tasting.

	NEIGHBORHOOD	HOUSES SOLD	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
15	INWOOD	50	Café	Mexican Restaurant	Park	Spanish Restaurant	American Restaurant	Wine Shop	Wine Bar	Bakery	Frozen Yogurt Shop	Chinese Restaurant
22	MIDTOWN WEST	746	Mexican Restaurant	Thai Restaurant	Italian Restaurant	Coffee Shop	Theater	Wine Shop	Gym	Wine Bar	Gourmet Shop	Pizza Place
23	MORNINGSIDE HEIGHTS	2	Deli / Bodega	Sandwich Place	Coffee Shop	Mexican Restaurant	Italian Restaurant	Café	Chinese Restaurant	Pharmacy	Park	College Cafeteria
30	WASHINGTON HEIGHTS	156	Pizza Place	Latin American Restaurant	Mexican Restaurant	Bookstore	Chinese Restaurant	Bakery	Coffee Shop	Thai Restaurant	Korean Restaurant	Park

RESULTS AND DISCUSSIONS



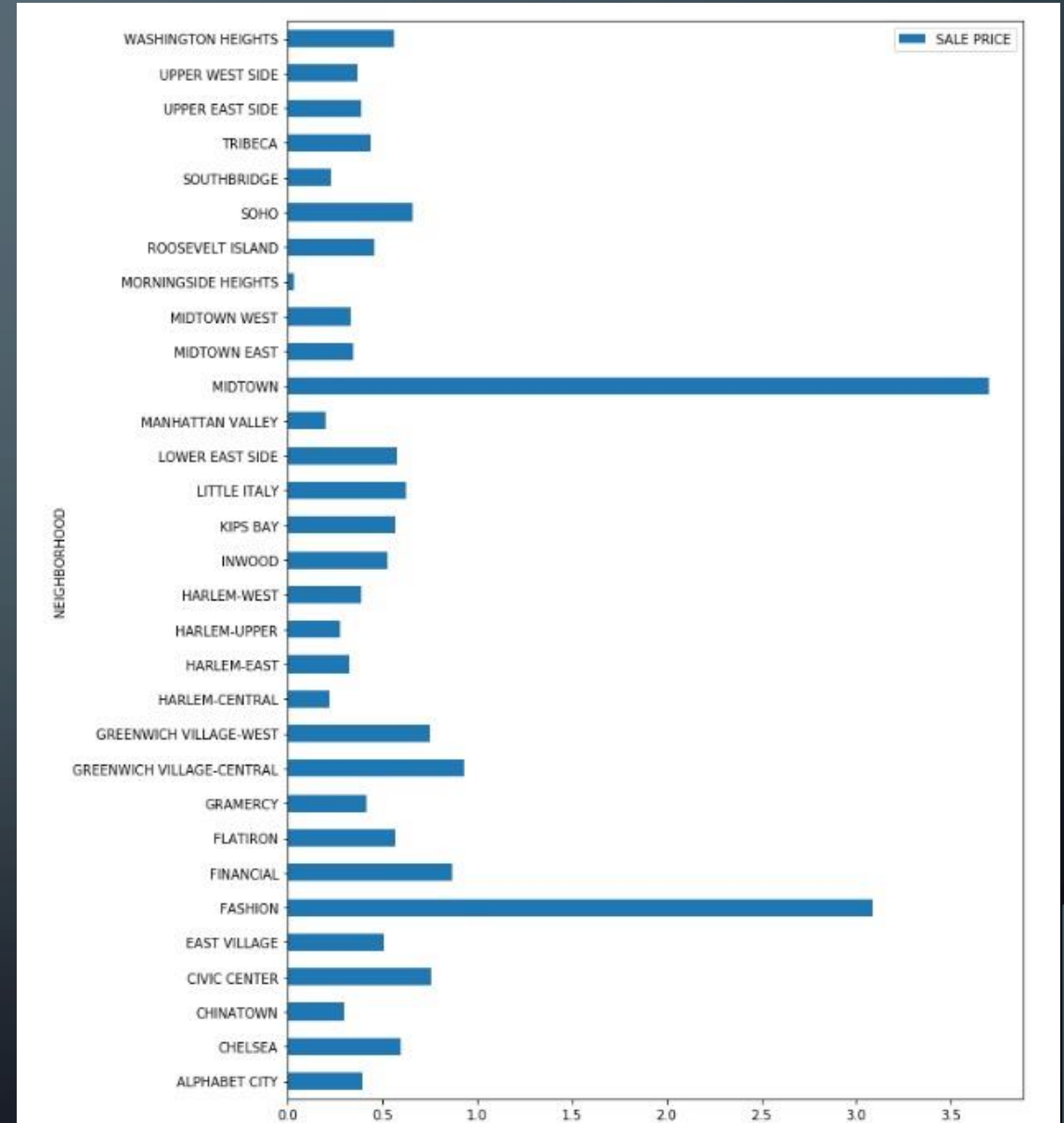
Additionally, it was possible to determine that the residents with the highest demand and probably the most accessible, correspond to those of **cluster 1**.

```
NEIGHBORHOOD
UPPER EAST SIDE    1387
UPPER WEST SIDE    1350
MIDTOWN WEST        746
HARLEM-CENTRAL      718
MIDTOWN EAST        671
Name: HOUSES SOLD, dtype: int64
```

RESULTS AND DISCUSSIONS

Also, the first 4 most expensive neighborhoods in the Manhattan borough belong to cluster 3, this may be due to the fact that it is the most visited area.

NEIGHBORHOOD	SALE PRICE
MIDTOWN	3.697168e+07
FASHION	3.087506e+07
GREENWICH VILLAGE-CENTRAL	9.282142e+06
FINANCIAL	8.673649e+06
CIVIC CENTER	7.594067e+06



CONCLUSION

- After collecting the information, grouping it and analyzing it, it was possible to categorize the different neighborhoods that make up the Manhattan borough, in this way the possible buyers or tenants will be able to consider and make the best decision when purchasing or renting an apartment. Choosing the neighborhood that best suits your tastes, interests or needs, depending on the places to go in the vicinity of the neighborhood.
- In addition, this could lead to great opportunities for the Manhattan real estate sector, it would contribute as an important resource in the final consideration or purchase stage, since it would add value to the property to be promoted.
- Also, it is important to highlight that in carrying out this type of modeling, an exhaustive review of the data entered is essential for its execution, since a minimum error could alter the results.