

Manhattan Real Estate Purchase
Price correlation to
Neighborhoods and surrounding
Venues

Elnatan Mataev

Problem statement

Main question:

- Is there a correlation between the type and number of venues in a specific neighborhood in Manhattan to its residential real estate cost?

Sub questions – other factors that impact the cost of housing

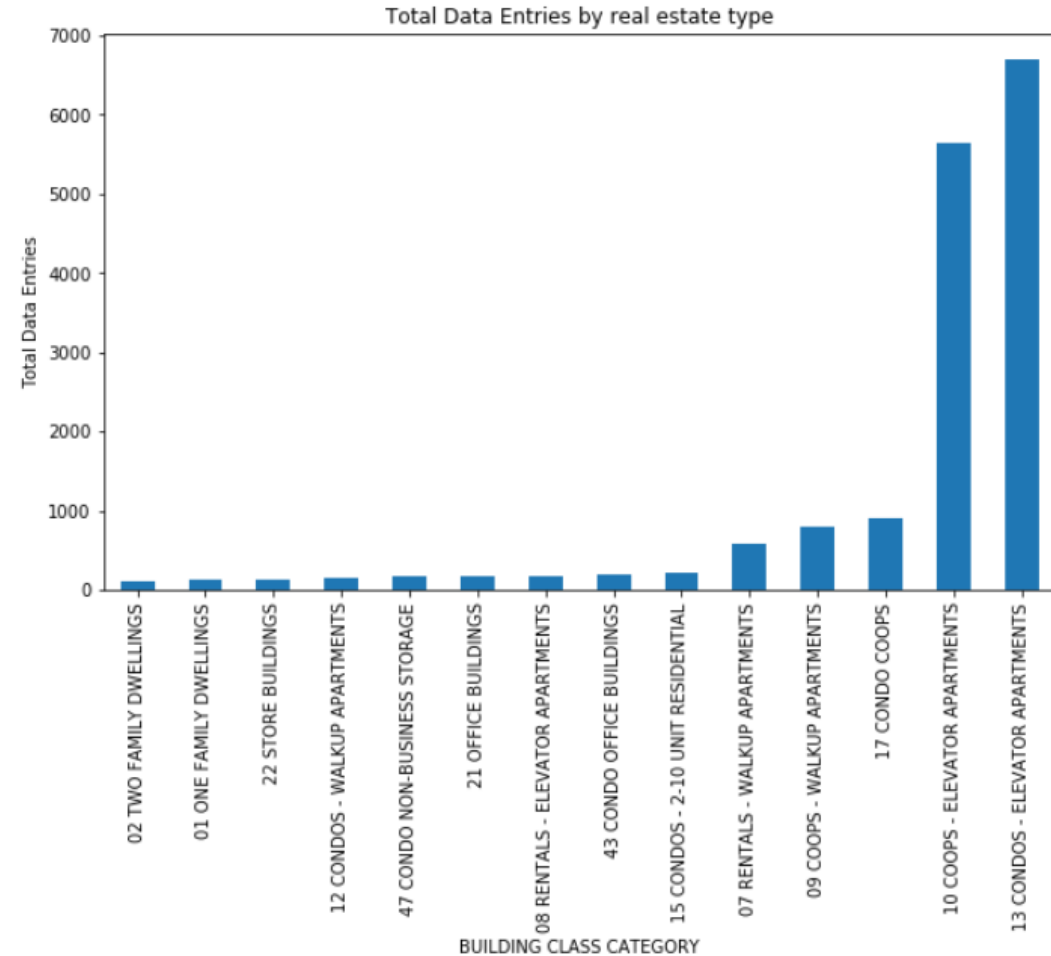
- Can we predict the cost of a residential unit in Manhattan based on its neighborhood?
- Impact of residential type on its price
- What is cost per square foot in Manhattan?
- Does an elevator make a difference to the cost of the residential unit? How much?

Dataset:

- Manhattan 2019 tax database on property sales tax (available online).
- Foursquare venues open-source
- Total of 16473 data entries → 16,473 transactions related to property purchases only in Manhattan on 2019 (Residential and commercial combined).

Total number of data entries by real estate building class category

- As anticipated – NYC has mostly apartment and co-op
- Very little one family or two families dwelling.



Example of data table as received from NYC real estate data base - 2019

	BOROUGH	NEIGHBORHOOD	BUILDING CLASS CATEGORY	TAX CLASS AT PRESENT	BLOCK	LOT	EASE-MENT	BUILDING CLASS AT PRESENT	ADDRESS	APARTMENT NUMBER	...	COMMERCIAL UNITS	TOTAL UNITS	LAND SQUARE FEET	GROSS SQUARE FEET	YEAR BUILT
0	1	ALPHABET CITY	01 ONE FAMILY DWELLINGS	1	376	43	NaN	S1	743 EAST 6TH STREET	NaN	...	1.0	2.0	2,090	3,680	1940.0
1	1	ALPHABET CITY	01 ONE FAMILY DWELLINGS	1	390	61	NaN	A4	189 EAST 7TH STREET	NaN	...	0.0	1.0	987	2,183	1860.0
2	1	ALPHABET CITY	02 TWO FAMILY DWELLINGS	1	404	1	NaN	B9	166 AVENUE A	NaN	...	0.0	2.0	1,510	4,520	1900.0
3	1	ALPHABET CITY	03 THREE FAMILY DWELLINGS	1	377	56	NaN	C0	263 EAST 7TH STREET	NaN	...	0.0	3.0	2,430	3,600	1899.0
4	1	ALPHABET CITY	03 THREE FAMILY DWELLINGS	1	393	9	NaN	C0	604 EAST 11TH STREET	NaN	...	0.0	3.0	2,375	5,110	1939.0

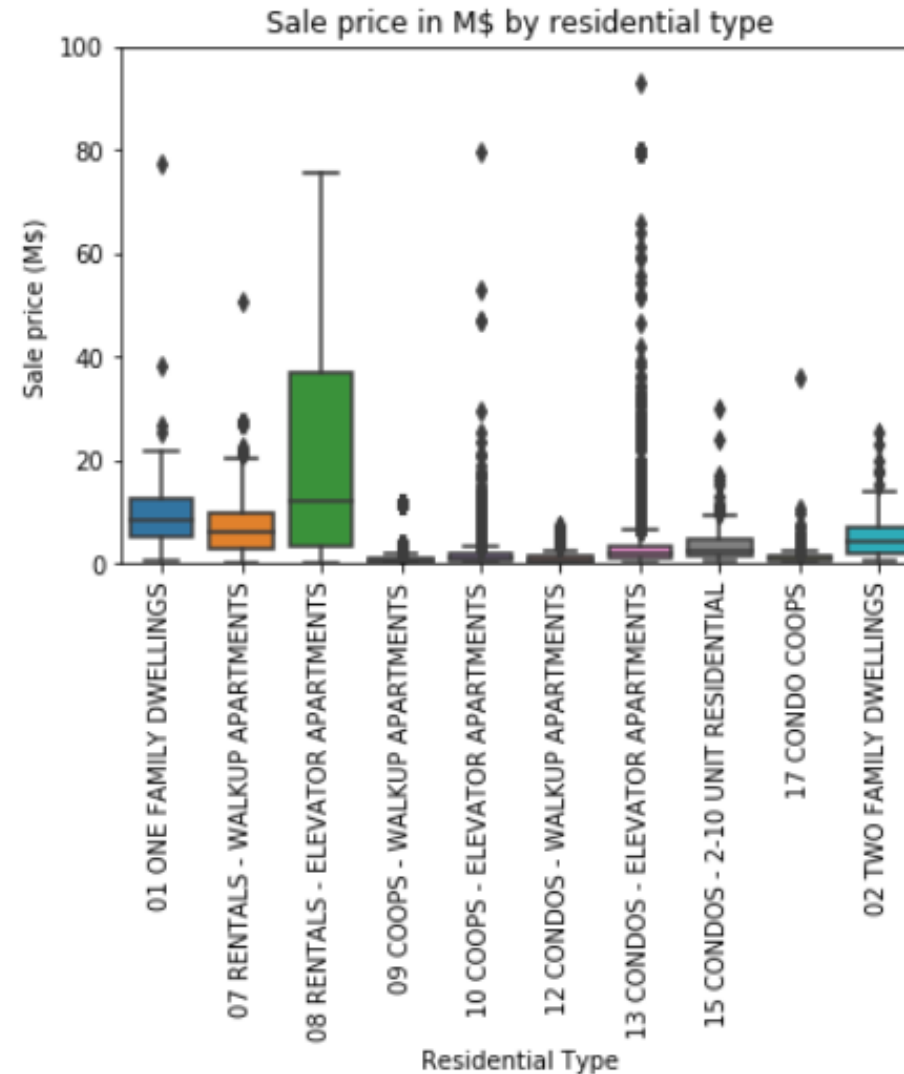
Real-estate data – post data cleanup

	BOROUGH	Hood	BUILDING CLASS CATEGORY	ZIP CODE	TOTAL UNITS	LAND SQUARE FEET	GROSS SQUARE FEET	YEAR BUILT	SALE DATE	SALE PRICE
0	1	East Village	01 ONE FAMILY DWELLINGS	10009	2.0	2,090	3,680	1940.0	7/24/2019	3,200,000
1	1	East Village	01 ONE FAMILY DWELLINGS	10009	1.0	987	2,183	1860.0	9/25/2019	0
2	1	East Village	02 TWO FAMILY DWELLINGS	10009	2.0	1,510	4,520	1900.0	7/22/2019	0
3	1	East Village	03 THREE FAMILY DWELLINGS	10009	3.0	2,430	3,600	1899.0	4/30/2019	6,300,000
4	1	East Village	03 THREE FAMILY DWELLINGS	10009	3.0	2,375	5,110	1939.0	10/24/2019	0

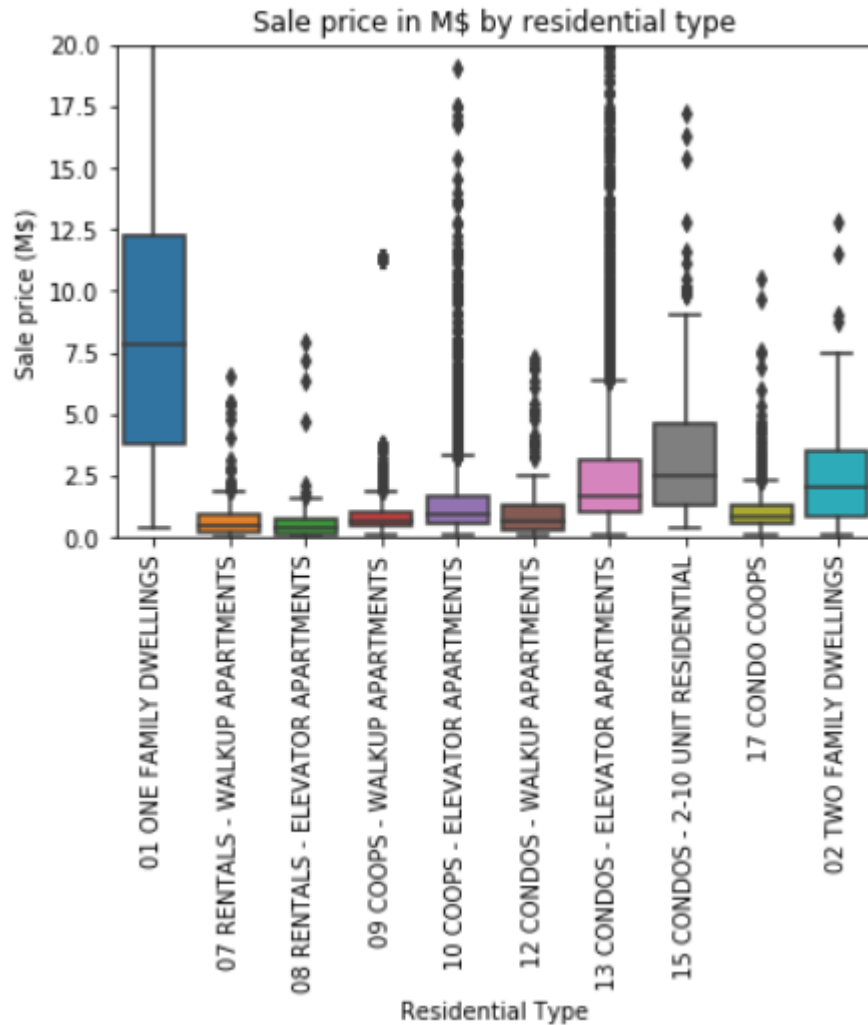
- In this work- neighborhood, sale price, building class category, total unit and gross square feet is taken into account
- Sale price =0 was dropped from the dataset.
- Only residential properties were taken into account. Commercial property sales were not considered.

Sale price vs. residential type

- Rentals properties (properties for rental purposes shows the most expensive sale price.
- Reason – multi-units under the same building (same purchase)
- Solution – normalizing the sale prices by the number of units



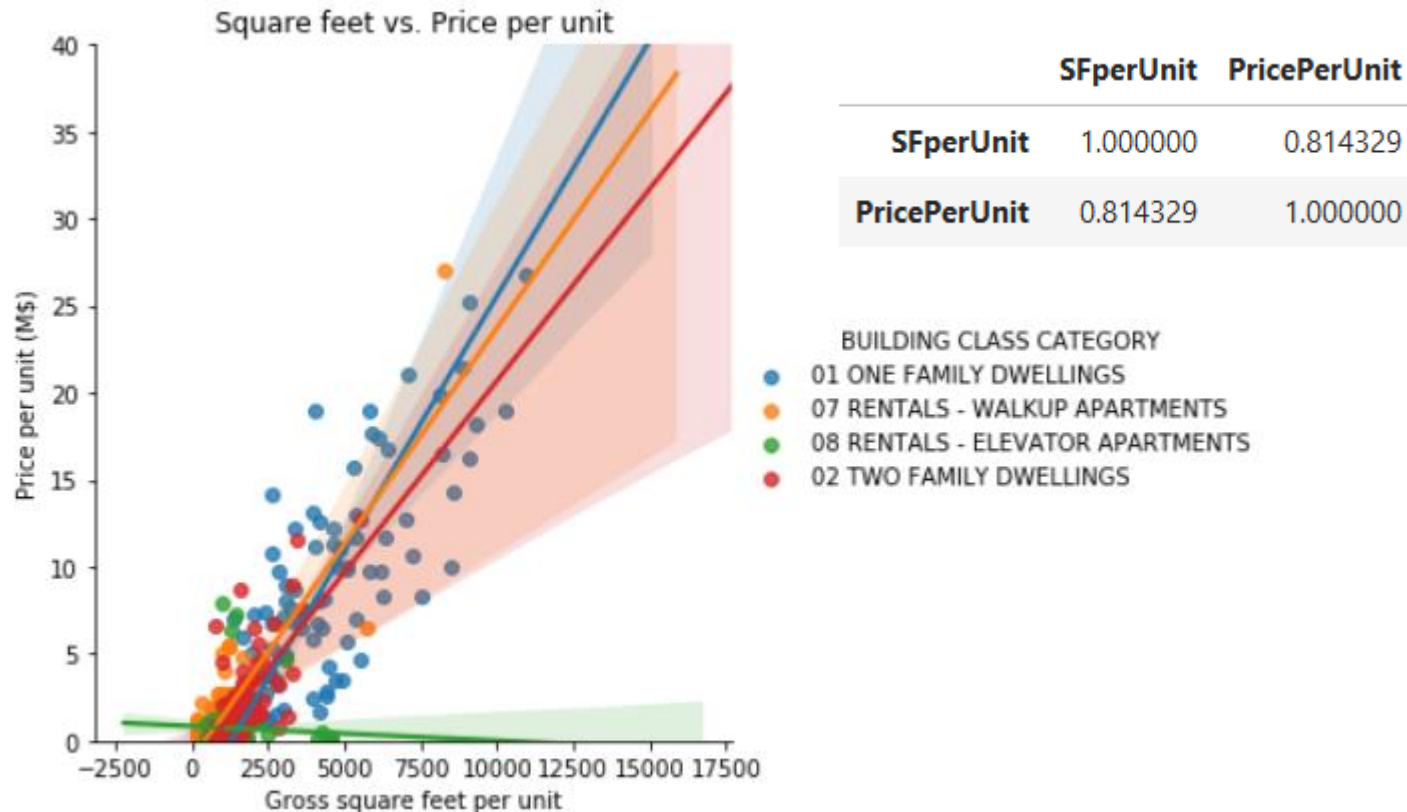
Sale price per unit vs. residential type



	BUILDING CLASS CATEGORY	PricePerUnit
0	01 ONE FAMILY DWELLINGS	7.825000
1	02 TWO FAMILY DWELLINGS	2.032500
2	07 RENTALS - WALKUP APARTMENTS	0.500000
3	08 RENTALS - ELEVATOR APARTMENTS	0.346698
4	09 COOPS - WALKUP APARTMENTS	0.625000
5	10 COOPS - ELEVATOR APARTMENTS	0.880000
6	12 CONDOS - WALKUP APARTMENTS	0.677500
7	13 CONDOS - ELEVATOR APARTMENTS	1.700000
8	15 CONDOS - 2-10 UNIT RESIDENTIAL	2.475000
9	17 CONDO COOPS	0.810000

- As anticipated – Single family and 2 family dwelling is the most expensive in NYC and not very common.
- Elevator condos and co-ops worth more than non-elevator. Adding roughly 200K\$ (co-op) -1M\$ (condo) to its value.
- Elevator did not add value to rental property

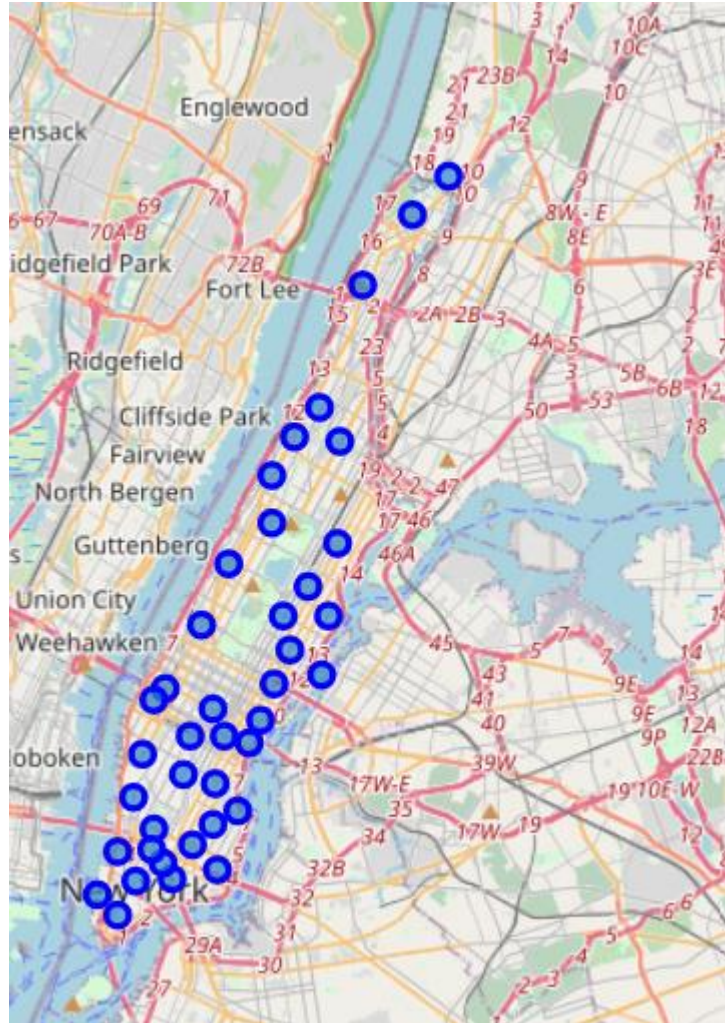
Gross square feet vs. real-estate price per unit



- Strong linear correlation between the size of the property to its price with $0.814 R^2$
- Slope = 0.002 in other words 1 square feet in Manhattan costs roughly 2000\$.
- **Very high** cost per square foot.

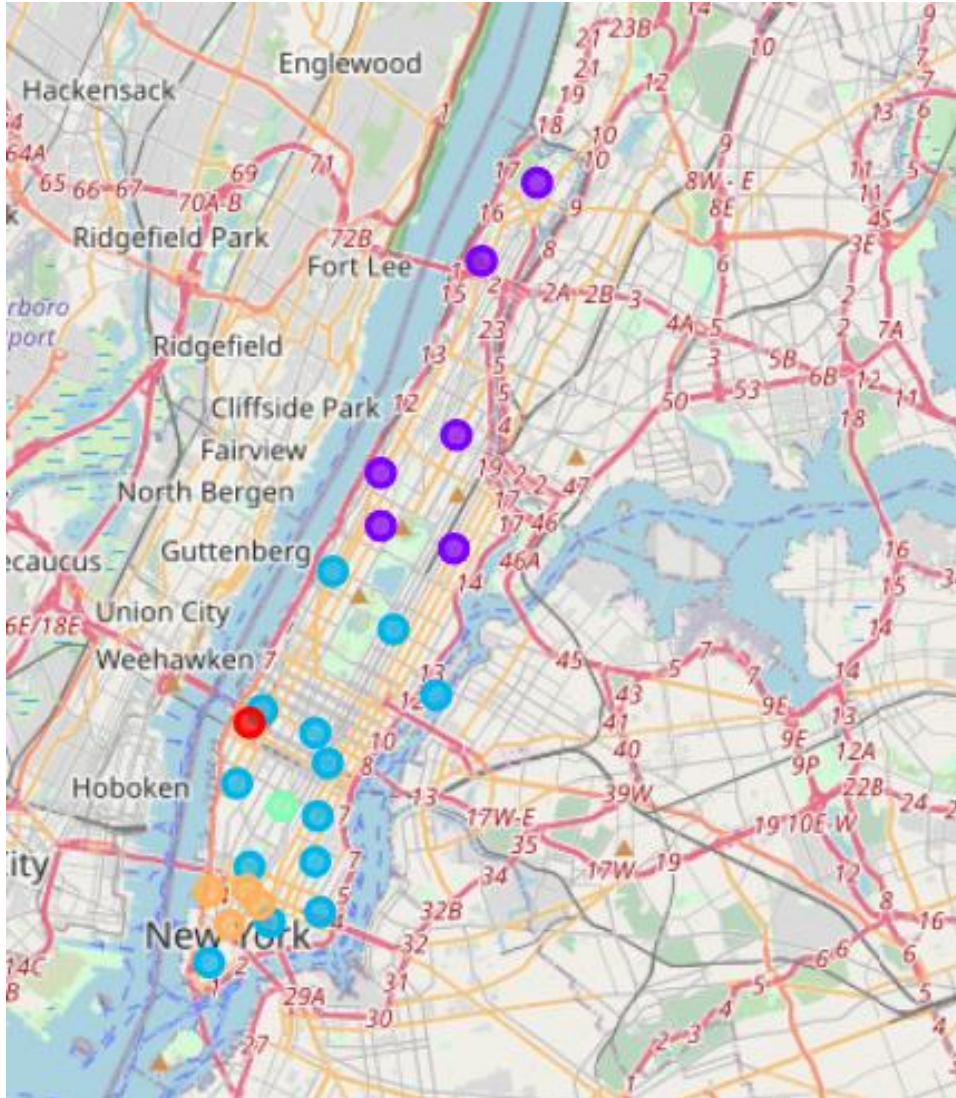
```
PricePerUnit=Slope*SFPerUnit + Intercept  
Intercept: [-1.49311099]  
Slope: [[0.00237466]]
```

Neighborhood locators based on Foursquare database



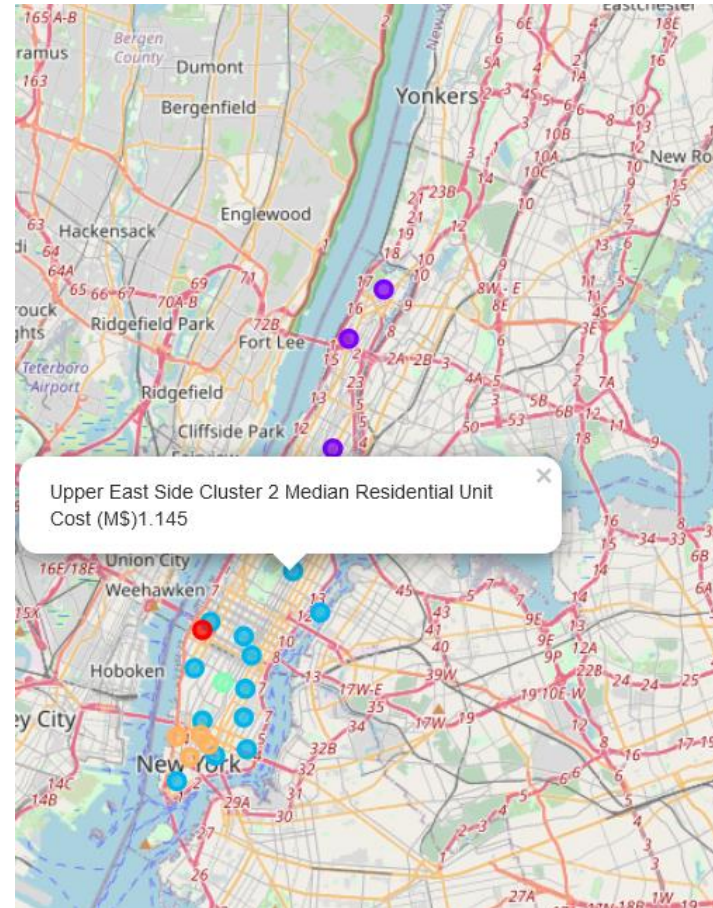
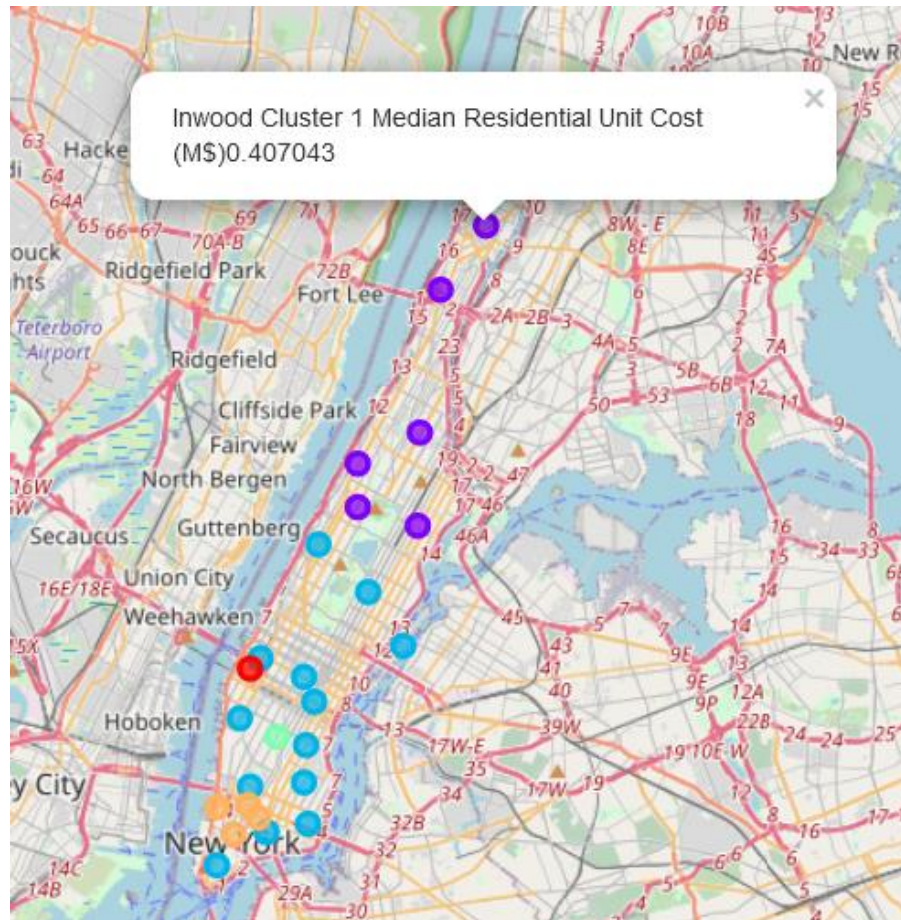
- Manhattan neighborhood latitude and longitude coordinates were taken from Foursquare open-source database.
- Neighborhood name was used as identifier.

Manhattan median real-estate purchase price by neighborhood



- Cluster 1 – Median price per unit < 0.75 M\$ (purple)
 - Cluster 2 – $0.75 < \text{Median price per unit} < 1.5$ M\$ (blue)
 - Cluster 3 – $1.5 < \text{Median price per unit} < 2.25$ M\$ (cyan)
 - Cluster 4 – $2.25 < \text{Median price per unit} < 3$ M\$ (orange)
 - Cluster 5 – Median price per unit > 3 M\$ (red)
-
- North Manhattan – more affordable housing – residence are typically working-class lower income families
 - Financial district area – expensive high-end housing. Typically residents work in the surrounding financial institutions → High income
 - Cluster 5- Hudson Yards –
 - Most expensive.
 - Recently renovated and developed.
 - Top notch shopping centers, parks and facilities
 - High-end houses with unique design

Manhattan median real-estate purchase price by neighborhood – interactive map



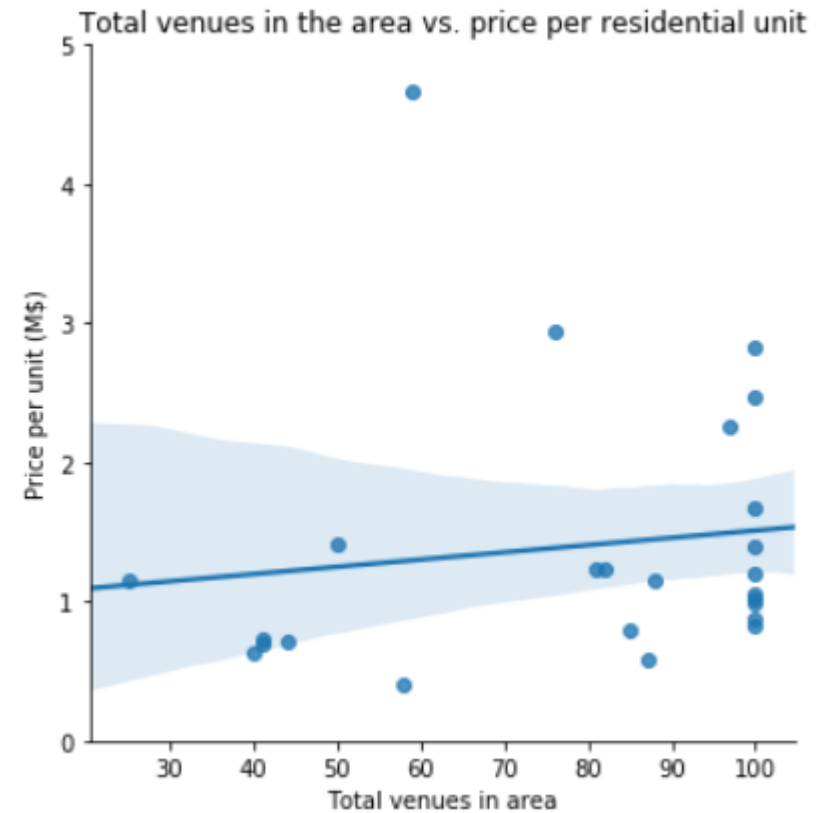
- 5 clusters by median price
- Popup shows the median real-estate price, and the cluster

Manhattan median real-estate purchase price by neighborhood – summary table

	Hood	PricePerUnit
0	Central Harlem	0.710000
1	Chelsea	1.195000
2	Chinatown	1.059500
3	Civic Center	2.466200
4	Clinton	0.871500
5	East Harlem	0.701500
6	East Village	0.820000
7	Financial District	1.022250
8	Flatiron	1.675000
9	Gramercy	1.225000
10	Greenwich Village	1.400000
11	Hudson Yards	4.666812
12	Inwood	0.407043
13	Little Italy	2.825000
14	Lower East Side	1.406496
15	Manhattan Valley	0.735000
16	Midtown	0.990000
17	Morningside Heights	0.627500
18	Murray Hill	0.787500
19	Roosevelt Island	1.150000
20	Soho	2.250000
21	Tribeca	2.940000
22	Upper East Side	1.145000
23	Upper West Side	1.225000
24	Washington Heights	0.590147

Correlation between total number of venues in neighborhood to its housing cost

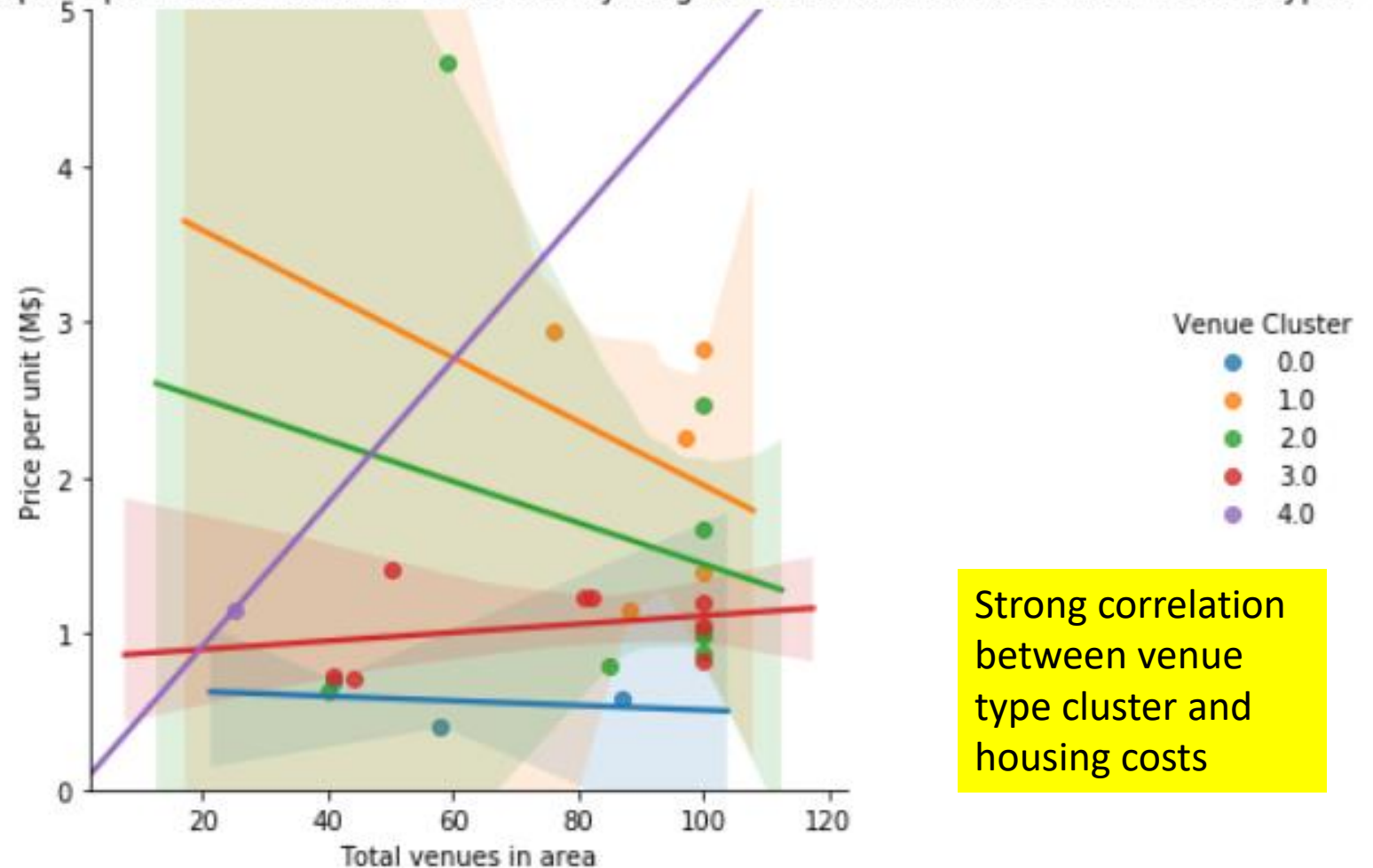
- Foursquare was used to extract the total amount of venues per neighborhood
- Slight correlation between the two factor.



Neighborhood **Venue Type** cluster vs. Housing Cost

- 2nd type of clustering in this work.
- Top 5 frequent venues in each neighborhood were considered and neighborhood were clustered accordingly.
- Clear separation in price vs. venue type cluster

Total venues in the area vs. price per residential unit - Clustered by neighborhood with most common venues types



Strong correlation between venue type cluster and housing costs

Neighborhood **Venue Type** cluster vs. Housing Cost

- **Venue type** cluster 0 → low end housing which are cheaper to buy.
 - Indeed the cluster 0 neighborhood are located in north Manhattan which has more affordable housing and typically has low-income families.
 - In the top 5 venues – every day food such as Mexican restaurants, Deli and bakeries.
- **Venue type** cluster 3 → Mid-range housing that are a little bit pricier than cluster 0.
 - All offer similar types of venues at similar price-point.
 - More mid-income venues such as bars, art galleries, desert shops, ice cream shops
- **Venue type** cluster 2 → Larger price variation across the neighborhoods .
 - If a buyer is interested in cluster 2 venues types, one can choose an apartment in a cheaper area that offers similar types of venues.
 - Venues include Wine-bars, hotels, coffee shops and fitness centers
- **Venue type** cluster 1 → High-end neighborhoods that is reflected in high-cost housing with a median cost of 2.11 M\$
 - Includes high end venues such as parks, comedy clubs, spas and sushi restaurants.

[249] :

	Venue Cluster	PricePerUnit
0	0.0	0.566230
1	1.0	2.112000
2	2.0	1.638345
3	3.0	1.046999
4	4.0	1.150000

Cluster 0

248]:

	Hood	PricePerUnit	Borough	Neighborhood_1	Latitude	Longitude	Price Cluster	Venue	Neighborhood_2	Venue Cluster	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
5	East Harlem	0.701500	Manhattan	East Harlem	40.792249	-73.944182	1	41	East Harlem	0.0	Mexican Restaurant	Bakery	Thai Restaurant	Deli / Bodega	Latin American Restaurant
12	Inwood	0.407043	Manhattan	Inwood	40.867684	-73.921210	1	58	Inwood	0.0	Mexican Restaurant	Lounge	Restaurant	Bakery	Café
24	Washington Heights	0.590147	Manhattan	Washington Heights	40.851903	-73.936900	1	87	Washington Heights	0.0	Café	Deli / Bodega	Bakery	Mobile Phone Shop	Latin American Restaurant

Let's check out the mean of price per unit by neighborhood cluster type

Cluster 1

[250] :

	Hood	PricePerUnit	Borough	Neighborhood_1	Latitude	Longitude	Price Cluster	Venue	Neighborhood_2	Venue Cluster	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
10	Greenwich Village	1.400	Manhattan	Greenwich Village	40.726933	-73.999914	2	100	Greenwich Village	1.0	Italian Restaurant	Café	Sushi Restaurant	Gym	Comedy Club
13	Little Italy	2.825	Manhattan	Little Italy	40.719324	-73.997305	4	100	Little Italy	1.0	Italian Restaurant	Bakery	Mediterranean Restaurant	Spa	Pizza Place
20	Soho	2.250	Manhattan	Soho	40.722184	-74.000657	4	97	Soho	1.0	Italian Restaurant	Sandwich Place	Mediterranean Restaurant	Clothing Store	Coffee Shop
21	Tribeca	2.940	Manhattan	Tribeca	40.721522	-74.010683	4	76	Tribeca	1.0	Italian Restaurant	American Restaurant	Park	Wine Bar	Greek Restaurant
22	Upper East Side	1.145	Manhattan	Upper East Side	40.775639	-73.960508	2	88	Upper East Side	1.0	Italian Restaurant	Coffee Shop	Bakery	Gym / Fitness Center	Yoga Studio

251]:

	Hood	PricePerUnit	Borough	Neighborhood_1	Latitude	Longitude	Price Cluster	Venue	Neighborhood_2	Venue Cluster	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
3	Civic Center	2.466200	Manhattan	Civic Center	40.715229	-74.005415	4	100	Civic Center	2.0	Coffee Shop	Hotel	Cocktail Bar	Spa	French Restaurant
4	Clinton	0.871500	Manhattan	Clinton	40.759101	-73.996119	2	100	Clinton	2.0	Theater	Coffee Shop	Gym / Fitness Center	Italian Restaurant	Wine Shop
7	Financial District	1.022250	Manhattan	Financial District	40.707107	-74.010665	2	100	Financial District	2.0	Coffee Shop	American Restaurant	Pizza Place	Café	Italian Restaurant
8	Flatiron	1.675000	Manhattan	Flatiron	40.739673	-73.990947	3	100	Flatiron	2.0	Gym / Fitness Center	Café	Italian Restaurant	Mediterranean Restaurant	Gym
11	Hudson Yards	4.666812	Manhattan	Hudson Yards	40.756658	-74.000111	5	59	Hudson Yards	2.0	Hotel	Italian Restaurant	Gym / Fitness Center	American Restaurant	Coffee Shop
16	Midtown	0.990000	Manhattan	Midtown	40.754691	-73.981669	2	100	Midtown	2.0	Hotel	Coffee Shop	Bakery	Theater	Cuban Restaurant
17	Morningside Heights	0.627500	Manhattan	Morningside Heights	40.808000	-73.963896	1	40	Morningside Heights	2.0	Park	American Restaurant	Coffee Shop	Bookstore	Sandwich Place
18	Murray Hill	0.787500	Manhattan	Murray Hill	40.748303	-73.978332	2	85	Murray Hill	2.0	Hotel	Sandwich Place	Coffee Shop	Gym / Fitness Center	Japanese Restaurant

Cluster 3

[252] :

	Hood	PricePerUnit	Borough	Neighborhood_1	Latitude	Longitude	Price Cluster	Venue	Neighborhood_2	Venue Cluster	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Central Harlem	0.710000	Manhattan	Central Harlem	40.815976	-73.943211	1	44	Central Harlem	3.0	African Restaurant	Chinese Restaurant	French Restaurant	Gym / Fitness Center	Cosmetics Shop
1	Chelsea	1.195000	Manhattan	Chelsea	40.744035	-74.003116	2	100	Chelsea	3.0	Coffee Shop	Art Gallery	Ice Cream Shop	Café	Bakery
2	Chinatown	1.059500	Manhattan	Chinatown	40.715618	-73.994279	2	100	Chinatown	3.0	Chinese Restaurant	Bakery	Cocktail Bar	Optical Shop	Spa
6	East Village	0.820000	Manhattan	East Village	40.727847	-73.982226	2	100	East Village	3.0	Bar	Mexican Restaurant	Coffee Shop	Cocktail Bar	Pizza Place
9	Gramercy	1.225000	Manhattan	Gramercy	40.737210	-73.981376	2	82	Gramercy	3.0	Bagel Shop	Coffee Shop	Bar	Pizza Place	American Restaurant
14	Lower East Side	1.406496	Manhattan	Lower East Side	40.717807	-73.980890	2	50	Lower East Side	3.0	Chinese Restaurant	Art Gallery	Pharmacy	Café	Cocktail Bar
15	Manhattan Valley	0.735000	Manhattan	Manhattan Valley	40.797307	-73.964286	1	41	Manhattan Valley	3.0	Coffee Shop	Yoga Studio	Mexican Restaurant	Bar	Pizza Place
23	Upper West Side	1.225000	Manhattan	Upper West Side	40.787658	-73.977059	2	81	Upper West Side	3.0	Italian Restaurant	Bar	Dessert Shop	Indian Restaurant	Wine Bar

Cluster 4

[253] :

	Hood	PricePerUnit	Borough	Neighborhood_1	Latitude	Longitude	Price Cluster	Venue	Neighborhood_2	Venue Cluster	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
19	Roosevelt Island	1.15	Manhattan	Roosevelt Island	40.76216	-73.949168	2	25	Roosevelt Island	4.0	Park	Plaza	School	Gym	Greek Restaurant