Analysis of the influence of nearby common venues on the square meter price of properties in Medellín City.

By Santiago Vélez

Different factors affect the price of a real estate property offered on the market

- In a property, both its own characteristics such as its finishes, distribution of the area or the size in square meters, as well as its external characteristics such as its geographical location in a city, have a direct influence on its commercial value.
- In this study, through a two-step analysis, it will be evaluated whether the common venues near properties influence or not their value. Expressed in the price per square meter in the city of Medellín Colombia.
- The findings can be used both for construction companies who want to start projects and for buyers and sellers of properties who want to know more about the common venues near the properties that interest them.

Data Section

Sources of data:

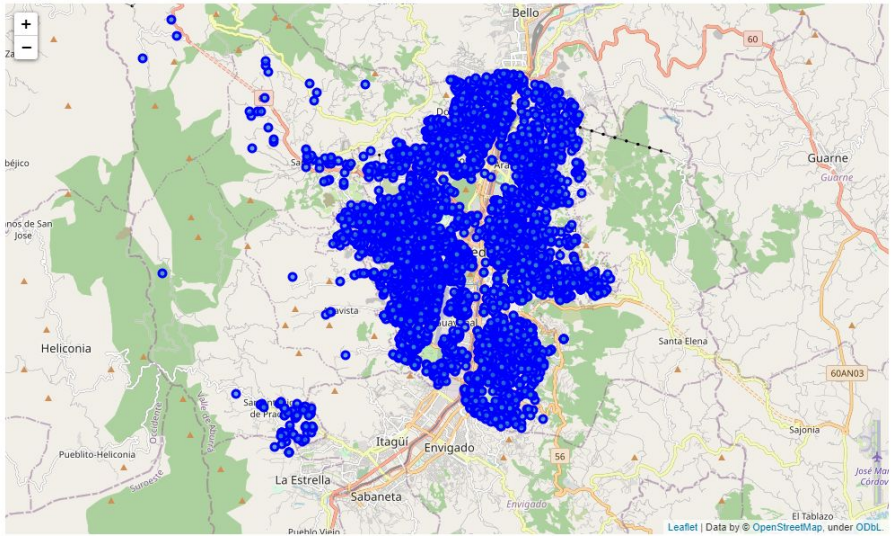
1. Real estate property data with its geographic location in Medellín, Colombia. (Aprox. 6100 records of properties in the dataset).
   Each property record includes:

   a. An unique code.
   b. Type of offer: (Sale or lease)
   c. Commercial Value
   d. Square meter value
   e. Geographical Coordinates.
2. Foursquare API to query common venues near previously classified real estate properties.

Methodology

Grouping of properties according to their value per square meter.
The complete property dataset consists of 15,725 records. When the filter is applied to get the properties for sale, the data is reduced to 6149 records.

The image shows the location of the properties on a map:



Methodology

Interquartile Analysis
Interquartile function over the ***Mt2 value*** variable

Interquartile group	Range in Mts2 Value	Description
Q1	0 - 900.000 COP	First 25% of properties
Q2	900000 - 1’412.000 COP	Second 25% of properties
Q3	1’412.000 - 2’000.000 COP	Third 25% of properties
Q4	> than 2’000.000 COP	Last 25% of properties

Methodology

Interquartile Analysis

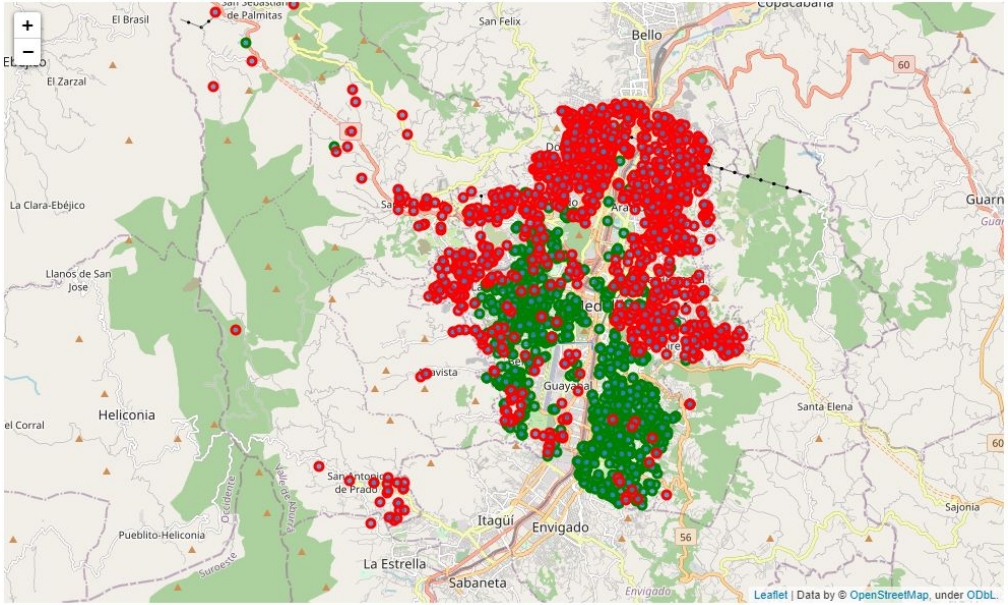
Using interquartile group as grouping criteria, a new column ***Quartile*** with the interquartile group can be added to the data frame.

	Property Code	Offer Type	Comercial Value	Mt2 Value	Coordinates	Quartile
2	76837	1	90000000	1034500.0	[-75.55950273469212, 6.281681757236898]	Q2
4	81000	1	100000000	1064000.0	[-75.55848773560459, 6.281669343283677]	Q2
7	91380	1	25000000	182500.0	[-75.55812114476444, 6.281473115285412]	Q1
13	81003	1	85000000	552000.0	[-75.56071377616377, 6.283044125616279]	Q1
16	79042	1	44430000	1234000.0	[-75.56041420318351, 6.283232305993462]	Q2

Methodology

Interquartile Groups Q1 and Q4
To contrast this two groups, this map is presented:

Q1: 1540 rows in RED **R**
Q4: 1488 rows in GREEN **G**



Methodology

Groups Q1 and Q4

Analysis of a random sample of properties for each Group.

For each property of the samples of groups Q1 and Q4, the 10 most common venues categories consulted and calculated in the Foursquare API are obtained.

Specifically, for this calculation, 100 venues within a radius of 200 meters are consulted for each property.

Properties of Q1

	Property Code	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	80374	Construction & Landscaping	Tram Station	Fish & Chips Shop	Bar	Business Service	Convenience Store	Dessert Shop	Fast Food Restaurant	Food Court	Tourist Information Center
1	80619	Fish & Chips Shop	Bar	Public Art	Dessert Shop	Tram Station	Business Service	Construction & Landscaping	Convenience Store	Fast Food Restaurant	Food Court
2	81013	Argentinian Restaurant	Fish & Chips Shop	Bar	Business Service	Construction & Landscaping	Convenience Store	Dessert Shop	Fast Food Restaurant	Tram Station	Tourist Information Center

Properties of Q4

	Property Code	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	76701	Clothing Store	Sandwich Place	BBQ Joint	Ice Cream Shop	Café	Karaoke Bar	Mexican Restaurant	Men's Store	Chinese Restaurant	Lounge
1	76784	Bakery	Yoga Studio	Eye Doctor	Colombian Restaurant	Creperie	Cupcake Shop	Dance Studio	Deli / Bodega	Diner	Donut Shop
2	77550	Park	Seafood Restaurant	Yoga Studio	Clothing Store	Coffee Shop	Colombian Restaurant	Creperie	Cupcake Shop	Dance Studio	Deli / Bodega

Methodology

Segmentation and clustering with k-means Machine learning algorithm

From the data of the common venues, we can submit each group of properties (Q1 and Q4) independently, to a K-means analysis where it will be possible to observe if the algorithm finds coincidences that tend to locate the 20 properties in the same large group.

For each group Q1 and Q4, a k-means analysis is performed with a value of k = 5 to obtain 5 groups (clusters) of properties.

Methodology

Segmentation and clustering with k-means Machine learning algorithm

Clusters label are added to the properties dataframe

Property Code	Offer Type	Comercial Value	Mt2 Value	Coordinates	Quartile	Cluster Labels	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
83797	1	26950000	350000.0	[-75.55520288649677, 6.241200755152511]	Q1	1	Tram Station	Seafood Restaurant	Food Court	Scenic Lookout	Public Art	Park	Motorcycle Shop
80619	1	95000000	772500.0	[-75.57270724106684, 6.291104299740561]	Q1	1	Fish & Chips Shop	Bar	Public Art	Dessert Shop	Tram Station	Business Service	Construction & Landscaping
86053	1	110000000	709500.0	[-75.58476106705798, 6.283855322877113]	Q1	1	Business Service	Convenience Store	Tram Station	Fish & Chips Shop	Bar	Construction & Landscaping	Dessert Shop
80374	1	101881000	893500.0	[-75.55940233917987, 6.237203767377195]	Q1	3	Construction & Landscaping	Tram Station	Fish & Chips Shop	Bar	Business Service	Convenience Store	Dessert Shop
85148	1	52726800	712500.0	[-75.57573302158212, 6.287084439615694]	Q1	2	Fast Food Restaurant	Tram Station	Fish & Chips Shop	Bar	Business Service	Construction & Landscaping	Convenience Store

Results Section

It is observed that for the Q1 group, close to the properties prevail:

- Viewpoints (Scenic Lookout)
- Subway (Train) stations
- Some businesses like bars and construction & landscaping
- tourist sites that are due to locations in the city center.

Types of venues that are associated with the attendance of people due to construction, tourist and public transport activities.
Another cluster, shows a similar behavior with common venues like train stations and construccions.

It is observed that for the Q4 group, close to the properties prevail venues with a greater variety of commercial categories prevail.
We could find common venues like:

- Hotels.
- Restaurants (of all kinds).
- Healthcare venues.
- Yoga, spa and gyms.

It can be inferred that these venues are associated with activities such as eating out, personal care, and national and international business areas (hotels and restaurants).

Discussion

Recommendations for investors and shareholders

The recommendations for investors and agents in the construction sector is that if they want to build residential properties which they want to sell at a high value according to their price per square meter (2,000,000 COP or higher), they should look for venues categories close, such as those described for the clusters in Q4.

By the contractor, if an investor seeks to create properties with lower price values per square meter (Less than 900,000 COP), in case of using the property for other types of activities, the nearby venues that influence this low price should be those described in the properties of group Q1

Conclusion

A qualitative relationship has been found between the categories of nearby common venues and the value of the price per square meter of a real estate property in the city of Medellín.

However, in the variability of the analyzes taking other random samples it is evidenced that other underlying variables may be correlated with the price of the properties, variables that could be found in more in-depth analysis regarding other characteristics of the properties and their surroundings.