



# **CLUSTERING SANTIAGO'S NEIGHBORHOOD CHILE**

# Clustering Santiago's Neighborhoods is helpful for people in moving plans

---

- The search for the best place to stay could be based on the nearest it is to the University, to School, to Job or maybe Price or Quality life.
- Criteria when you start to Research Will vary regarding priorities.
- Clustering Neighborhoods in Santiago will help foreign people planning to move to:
  - ✓ Get the common characteristics of every group of Neighborhoods.
  - ✓ Plan and focus their search.

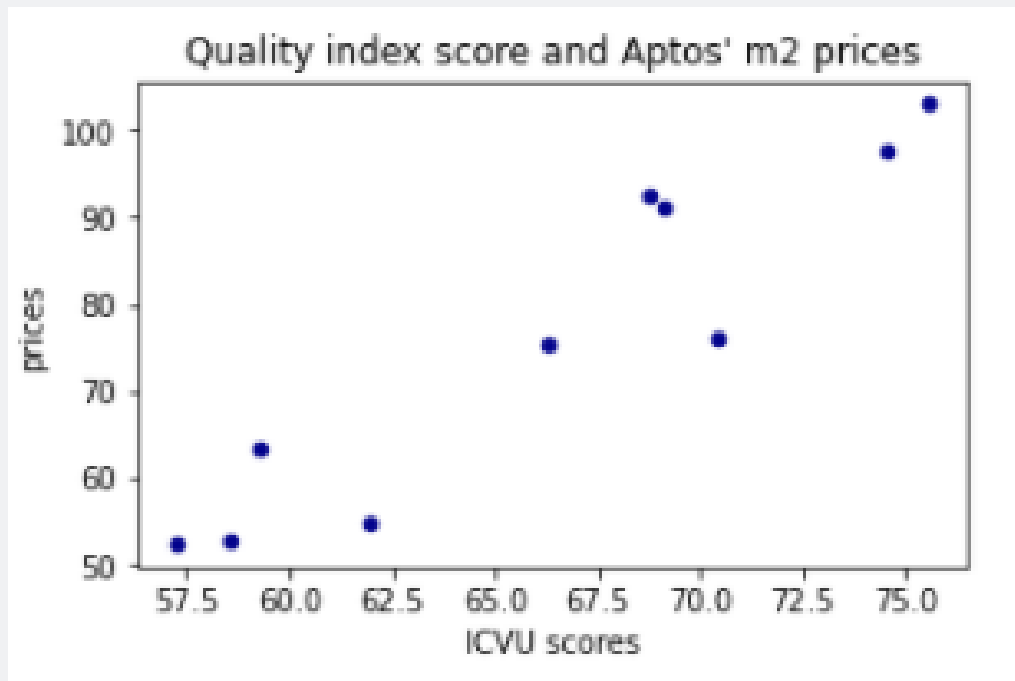
# Data Acquisition

---

- List of Neighborhoods scrapped from [Wikipedia](#)
- ICVU index (Urban Quality Life's Index) scrapped from [Emol's Newspaper](#).
- Apartments' m2 prices read from a CSV file collected manually and stored in the project file.
- Coordinates read from a CSV file collected manually and stored in the project file.
- Venues categories explored from Foursquare.
- All the information was merged based on the Top 10 Neighborhoods from Santiago in the ICVU ranking.
- Venues which categories were named "neighborhoods" were dropped (two rows). There are 791 rows in total and 187 Features (columns) to be included in the model (ICVU scores, m2 prices and 185 venues categories).

# Exploration

---



There is a relation between the ICVU scores and Apartments' m2 prices. It looks like the biggest the score is, the Price Will be higher.

# Exploration

|              | Neighborhood Latitude | Neighborhood Longitude | Venue | Venue Latitude | Venue Longitude | Venue Category | Pos. | Region | Province | Capital | latitude | longitude | ICVU Score | m2 price |
|--------------|-----------------------|------------------------|-------|----------------|-----------------|----------------|------|--------|----------|---------|----------|-----------|------------|----------|
| Neighborhood |                       |                        |       |                |                 |                |      |        |          |         |          |           |            |          |
| La Florida   | 24                    | 24                     | 24    | 24             | 24              | 24             | 24   | 24     | 24       | 24      | 24       | 24        | 24         | 24       |
| La Reina     | 60                    | 60                     | 60    | 60             | 60              | 60             | 60   | 60     | 60       | 60      | 60       | 60        | 60         | 60       |
| Las Condes   | 28                    | 28                     | 28    | 28             | 28              | 28             | 28   | 28     | 28       | 28      | 28       | 28        | 28         | 28       |
| Lo Barnechea | 100                   | 100                    | 100   | 100            | 100             | 100            | 100  | 100    | 100      | 100     | 100      | 100       | 100        | 100      |
| Macul        | 86                    | 86                     | 86    | 86             | 86              | 86             | 86   | 86     | 86       | 86      | 86       | 86        | 86         | 86       |
| Providencia  | 100                   | 100                    | 100   | 100            | 100             | 100            | 100  | 100    | 100      | 100     | 100      | 100       | 100        | 100      |
| San Miguel   | 94                    | 94                     | 94    | 94             | 94              | 94             | 94   | 94     | 94       | 94      | 94       | 94        | 94         | 94       |
| Santiago     | 99                    | 99                     | 99    | 99             | 99              | 99             | 99   | 99     | 99       | 99      | 99       | 99        | 99         | 99       |
| Vitacura     | 100                   | 100                    | 100   | 100            | 100             | 100            | 100  | 100    | 100      | 100     | 100      | 100       | 100        | 100      |
| Ñuñoa        | 100                   | 100                    | 100   | 100            | 100             | 100            | 100  | 100    | 100      | 100     | 100      | 100       | 100        | 100      |

Lo Barnechea, Providencia, Vitacura and Ñuñoa have the highest number of venues around. Las Condes and La Florida have the smaller number of venues around.

# Exploration

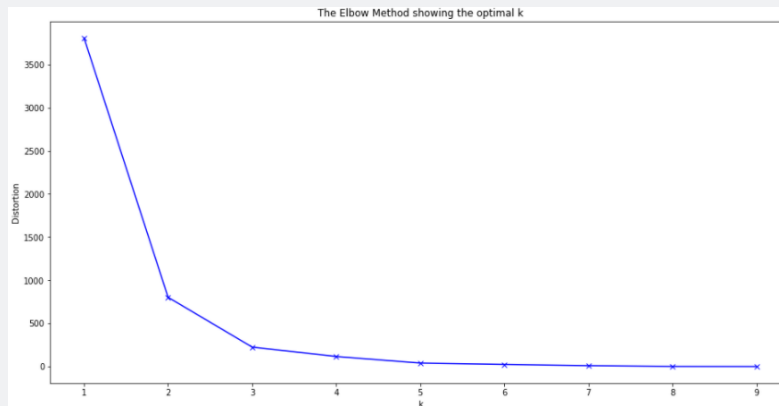
|   | Neighborhood | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue | 5th Most Common Venue |
|---|--------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 0 | La Florida   | Park                  | Gym                   | Restaurant            | Big Box Store         | Coffee Shop           |
| 1 | La Reina     | Restaurant            | Pizza Place           | Sushi Restaurant      | Chinese Restaurant    | Playground            |
| 2 | Las Condes   | Soccer Field          | Plaza                 | Park                  | Racetrack             | Pharmacy              |
| 3 | Lo Barnechea | Gym                   | Restaurant            | Coffee Shop           | Sushi Restaurant      | Pharmacy              |
| 4 | Macul        | Chinese Restaurant    | Restaurant            | Soccer Field          | Hot Dog Joint         | Sushi Restaurant      |
| 5 | Providencia  | Coffee Shop           | Park                  | Hotel                 | French Restaurant     | Ice Cream Shop        |
| 6 | San Miguel   | Sushi Restaurant      | Pizza Place           | Restaurant            | Plaza                 | Park                  |
| 7 | Santiago     | Coffee Shop           | Café                  | Pizza Place           | Hotel                 | Hostel                |
| 8 | Vitacura     | Café                  | Gym / Fitness Center  | Restaurant            | Bakery                | Italian Restaurant    |
| 9 | Ñuñoa        | Coffee Shop           | Bakery                | Plaza                 | Restaurant            | Pizza Place           |

The category most repeated in 1<sup>st</sup> place is Coffee (café) shop and then restaurants. Then we can see a mix of parks and gyms.

# Clustering model

K-means model is easy to apply and more efficient than other models, like Hierarchical.

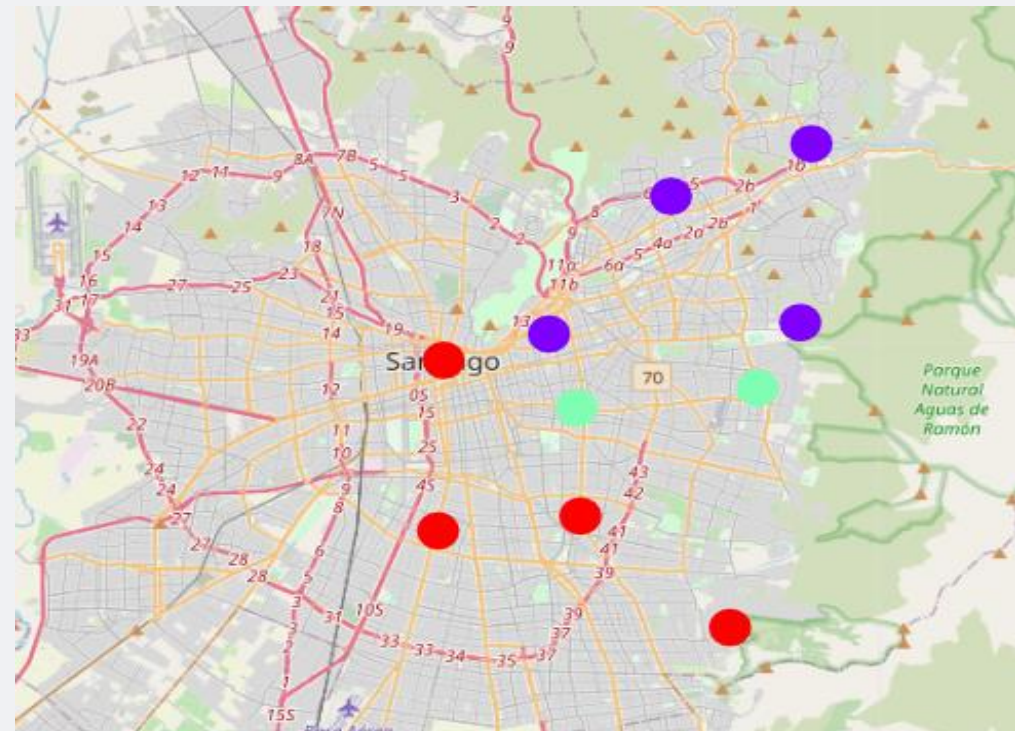
It iteratively works around two steps, reassigning centroids and every point of the data set to a new centroid until the error is minimized.



Elbow model.

$K = 3$

## Santiago's city



K-means algorithm

# Conclusion

---

- The model results in 3 clusters based mainly on the ICVU score and the apartments m2 prices. Even though the venues categories were included, there is not a clear differentiation between Neighborhoods with the information stored in Foursquare. But we can view that these have in common the restaurant categories, coffee shop, park and gyms.