



Photo by [Jorge Gardner](#) on [Unsplash](#)

Segmenting the best places to live in Bogota - Colombia

According to an economic, social and life
quality approach

Background

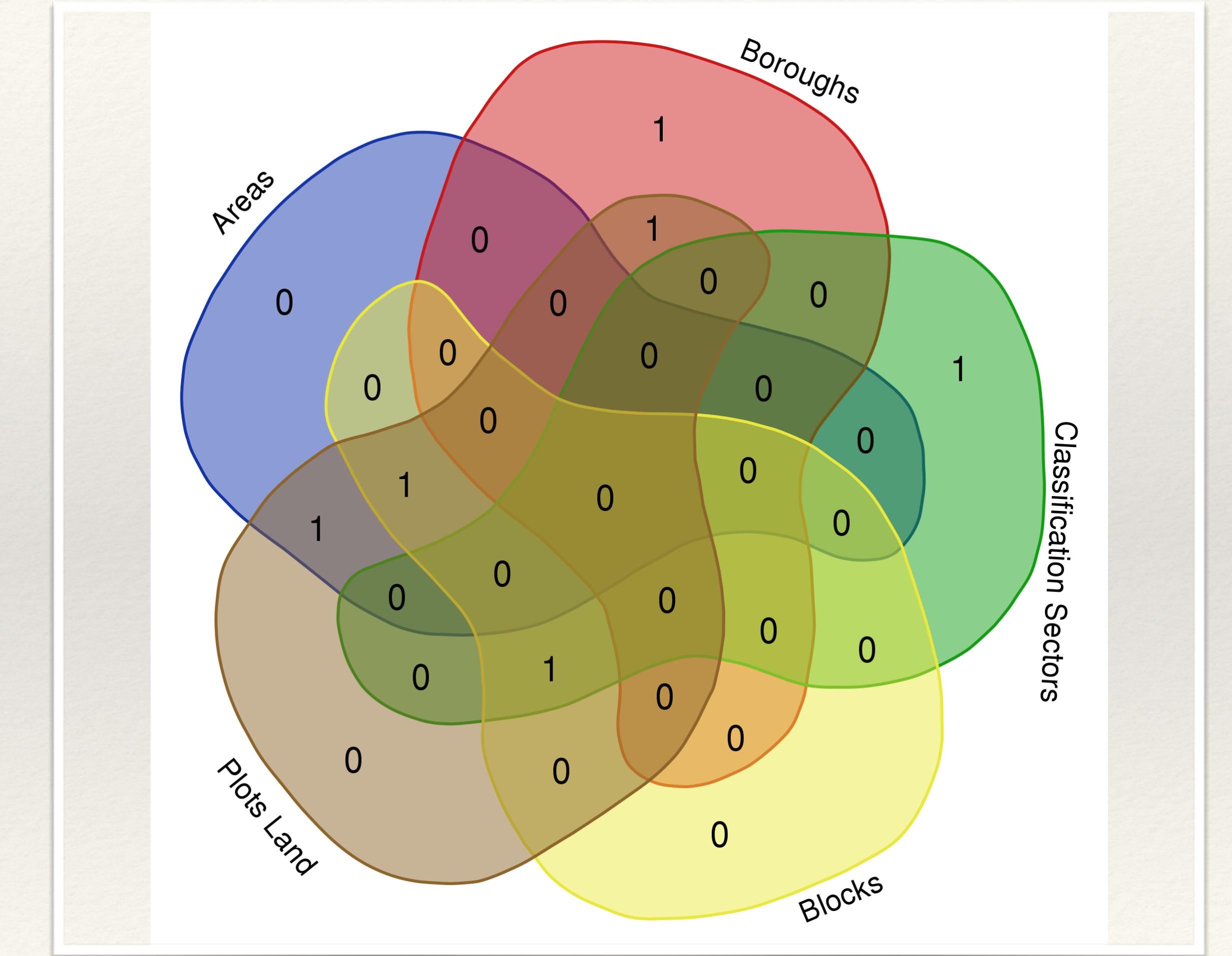
- ❖ Bogotá is the capital of Colombia and the administrative and economic center of the country. This city is divided into 20 boroughs, 112 zones planning units, 1169 cadastral areas and 1048 neighborhoods.
- ❖ Each area of the city is classified depending on the characteristics of the houses, the urban environment of the area, and the urban context. Thus, the city is subdivided into 6 socio-economic sectors, which the number 1 being the lowest and 6 the highest.
- ❖ The real estate market encompasses all these factors in order to provide a price and quality of life advantages that should be analyzed when choosing a home.

Business Problem

- ❖ A married couple lives in the suburbs of Bogota and every day they take 2 hours per way to arrive at home or work. Bogota has a big problem with its traffic, there are a lot of cars and public transportsations is not good at all as an incentive to take it. Because of this, they decide to move to Bogota under the following conditions:
 - ❖ The new place has to be close to their jobs. It is close to Parque de la 93
 - ❖ Only they can afford up to sector 4.
 - ❖ They do not want to live in sectors 1 and 2.
 - ❖ It should be a supermarket, pharmacies, and bus stations close to the new place.

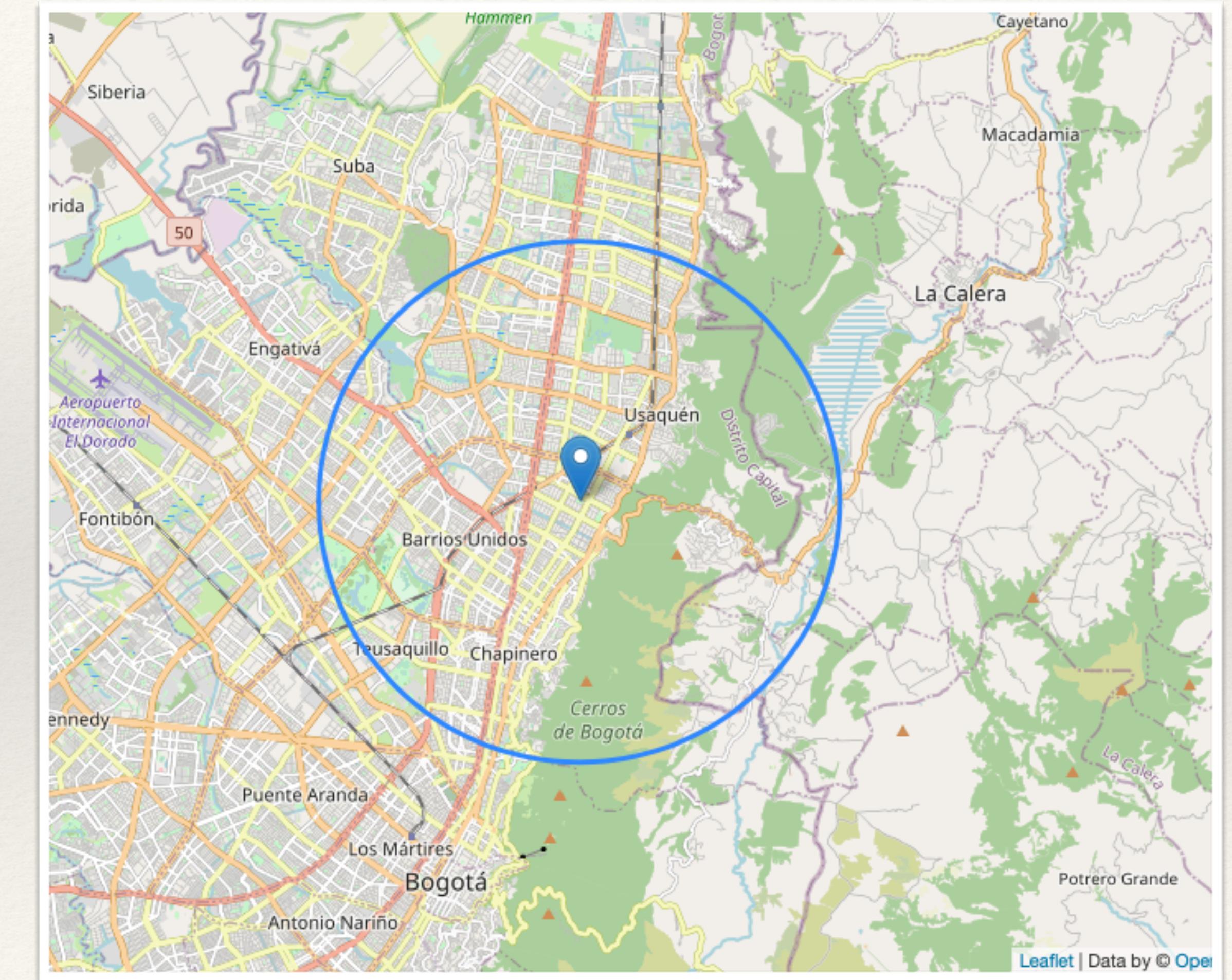
Data Acquisition and Cleaning

- ❖ Socioeconomic Datasets
 - ❖ From Colombia's Open Data Portal
 - ❖ Only keep codes, names from places and information geographical to show on a map.
- ❖ Venues Datasets
 - ❖ From Foursquare API
 - ❖ Using categories taken from a survey of the couple



Data selection

- ❖ The workplace is close to a venue site in Bogota called Parque de la 93 and will be established as the search centre with a radius of 6km.

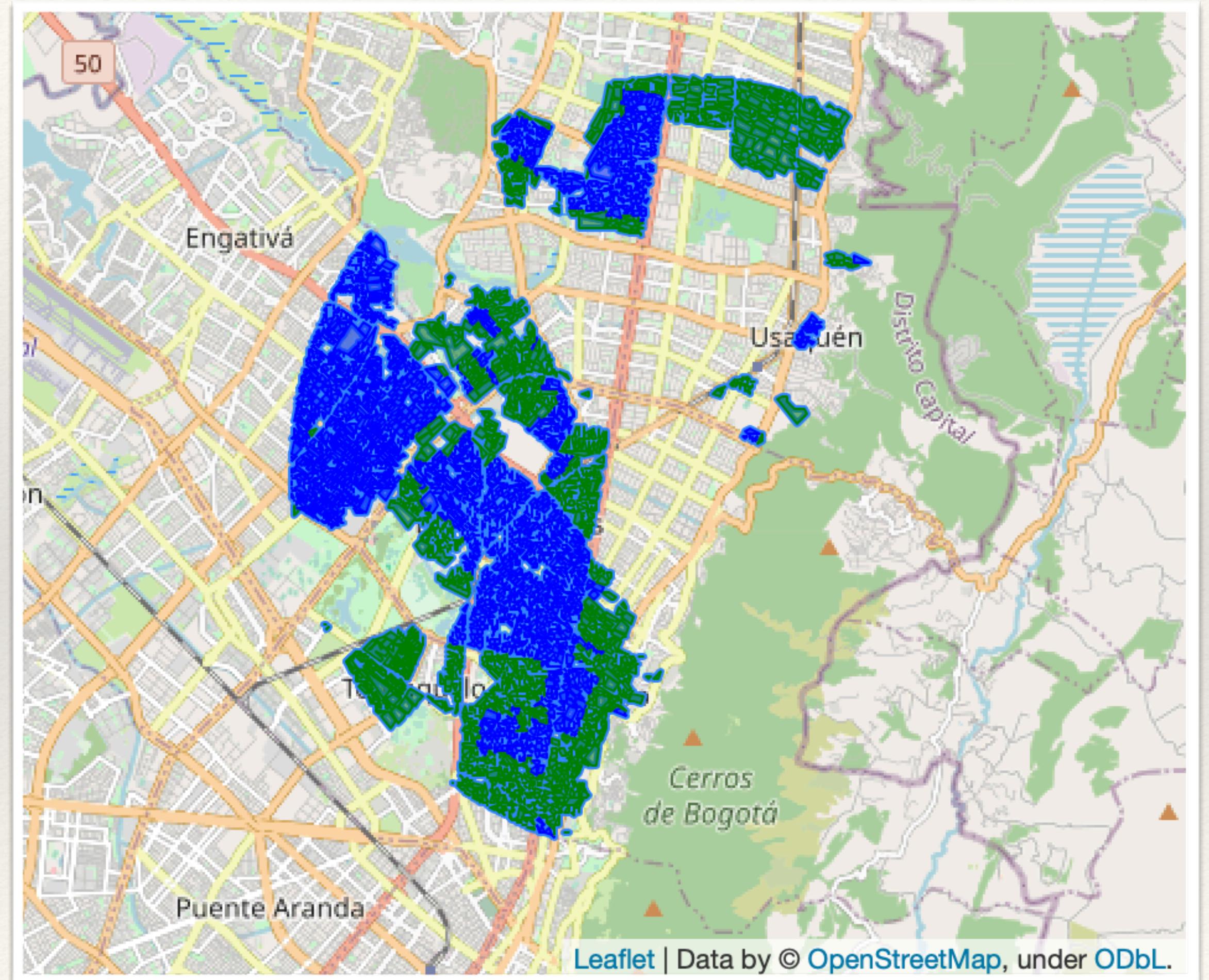


Methodology

- ❖ A radius of **6 km** will be selected.
- ❖ Only classification sectors 3 and 4 can be selected.
- ❖ Survey the stakeholders about what type of venues they would like to have near their future home.
- ❖ Foursquare database will be used to consult all of the venues that each center in each block has within a radius of 500 meters.
- ❖ Finally, a map will be presented highlighting each cluster created with the k-means technique to classify each point from its classification sectors and nearby locations as well as a small description of what can be found at each location.

Datasets Building

- ❖ For each point on the map, the venues of the selected categories are searched within a radius of 500 meters.
- ❖ Categories to use:
 - ❖ Arts Entertainment
 - ❖ Food Drink Shop
 - ❖ Health Services
 - ❖ NightLife
 - ❖ Outdoors
 - ❖ Professional Services
 - ❖ Public Transportation
 - ❖ Shop and Services
 - ❖ To Eat

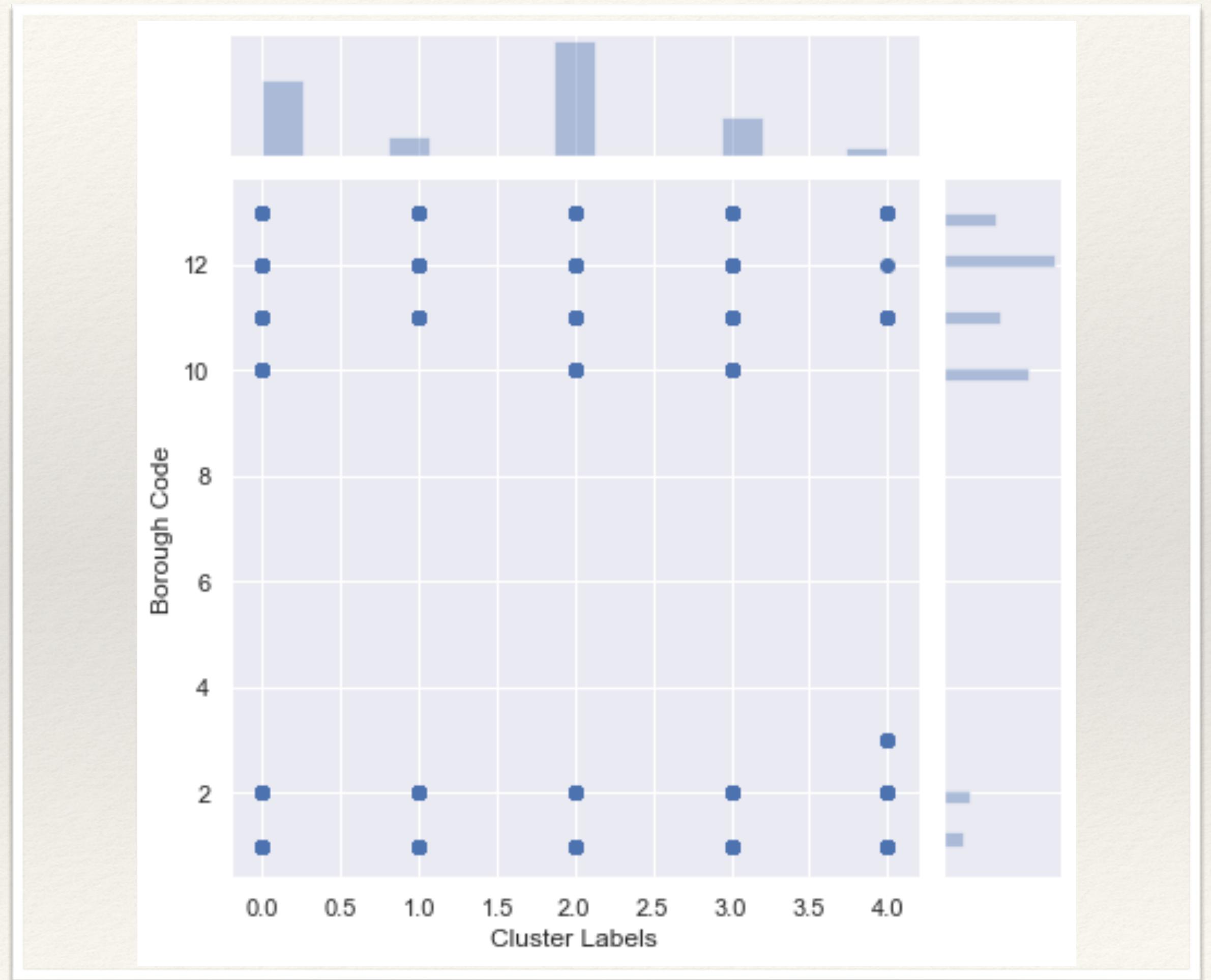


Clustering

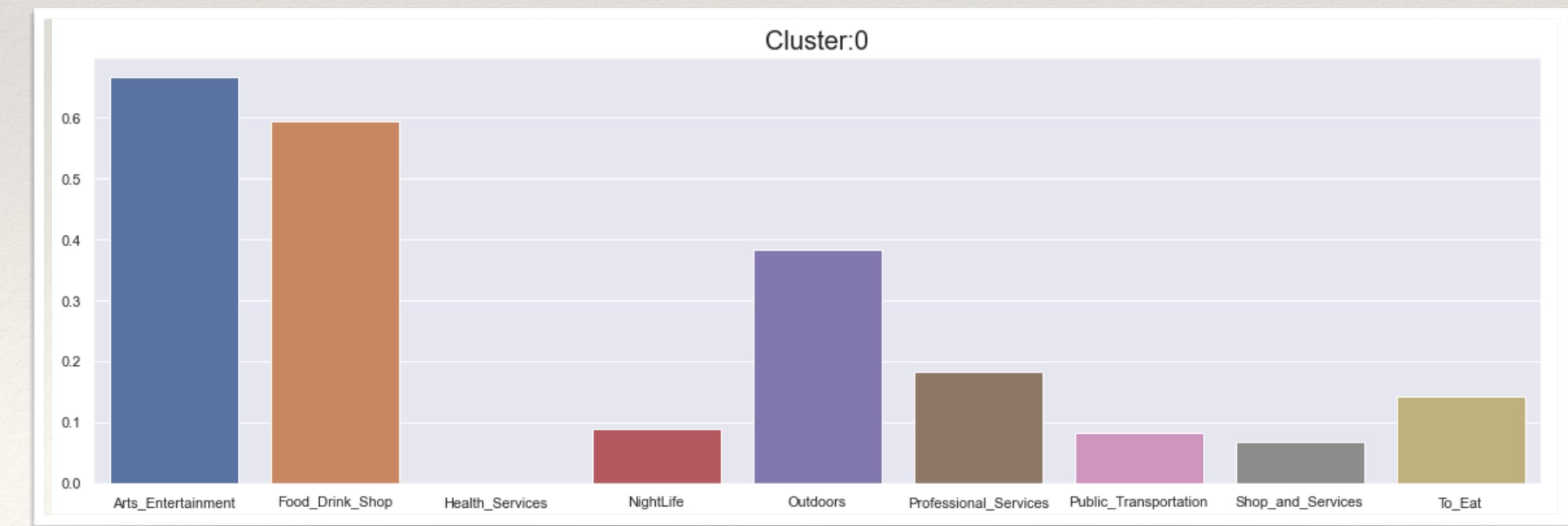
- ❖ 5 Cluster
 - ❖ K-Means Algorithm

| Cluster | Arts Entertainment | Food Drink Shop | Health Services | Night Life | Outdoors | Professional Services | Public Transportation | Shop and Services | To Eat |
|---------|--------------------|-----------------|-----------------|------------|----------|-----------------------|-----------------------|-------------------|----------|
| 0 | 0.666470 | 0.593852 | 0.0 | 0.088313 | 0.383231 | 0.183426 | 0.083138 | 0.066958 | 0.141254 |
| 1 | 1.000000 | 0.859930 | 0.0 | 0.745323 | 0.252241 | 0.686722 | 0.000000 | 0.904919 | 0.615705 |
| 2 | 0.000000 | 0.000000 | 0.0 | 0.000000 | 0.000000 | 0.000000 | 1.000000 | 0.000000 | 0.000000 |
| 3 | 0.537693 | 0.845650 | 0.0 | 0.263595 | 0.490323 | 0.558619 | 0.000000 | 0.417273 | 0.340280 |
| 4 | 0.649935 | 1.000000 | 0.0 | 1.000000 | 1.000000 | 1.000000 | 0.000000 | 1.000000 | 1.000000 |

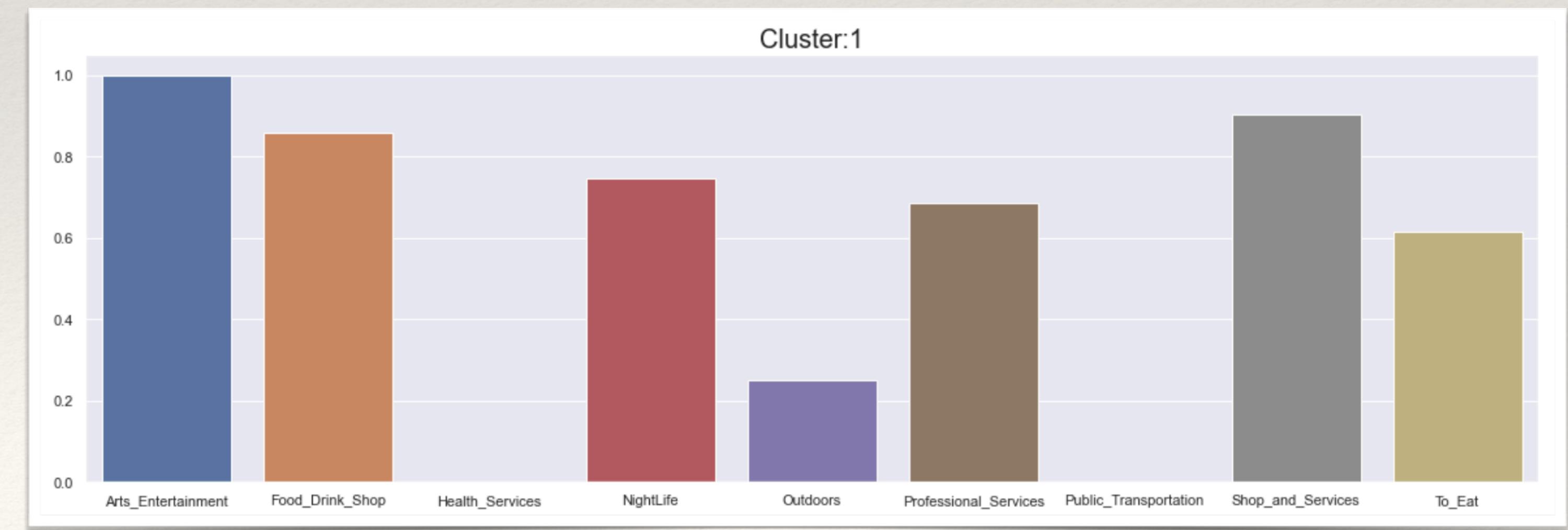
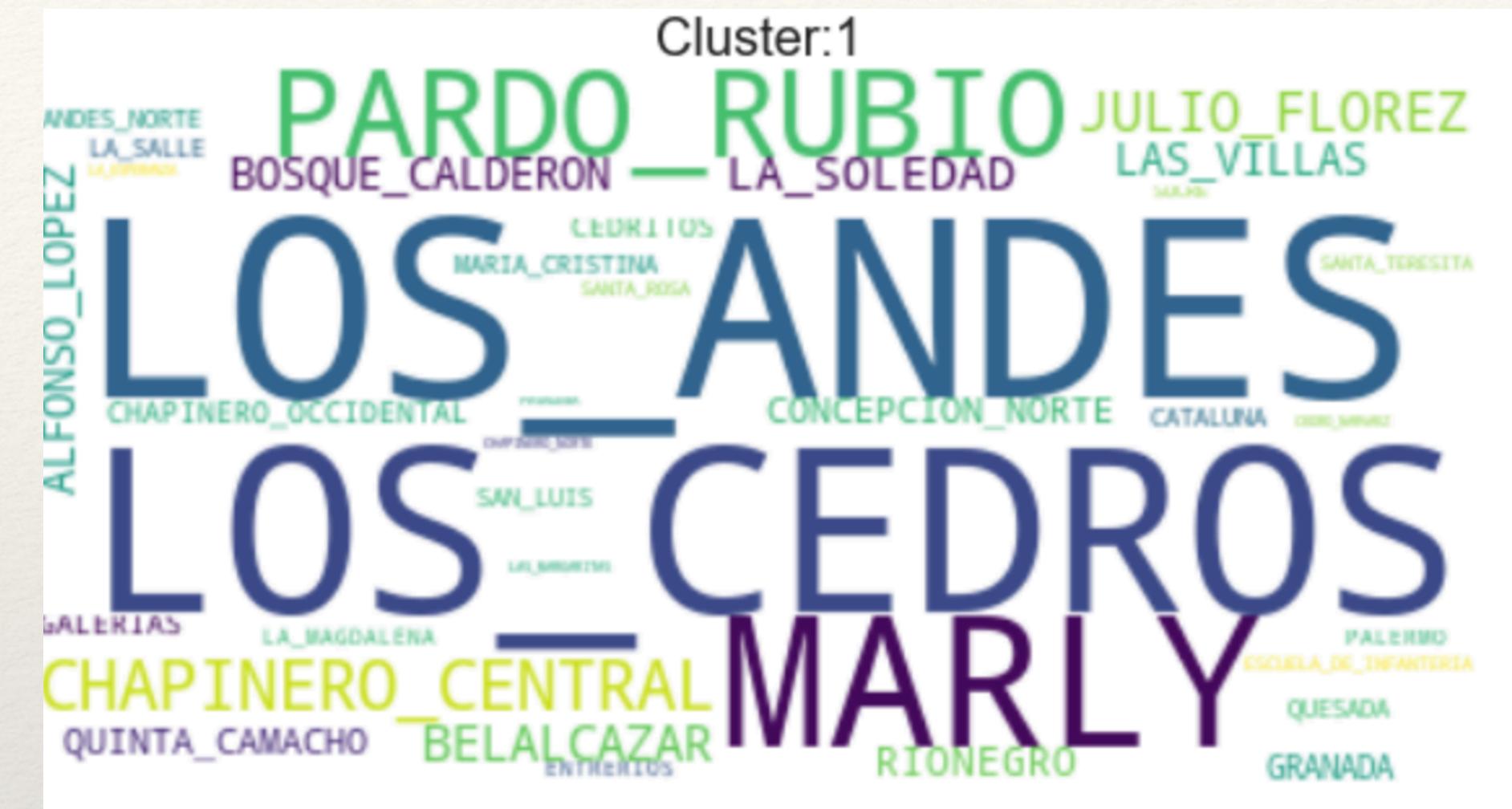
Normalized mean data of each cluster



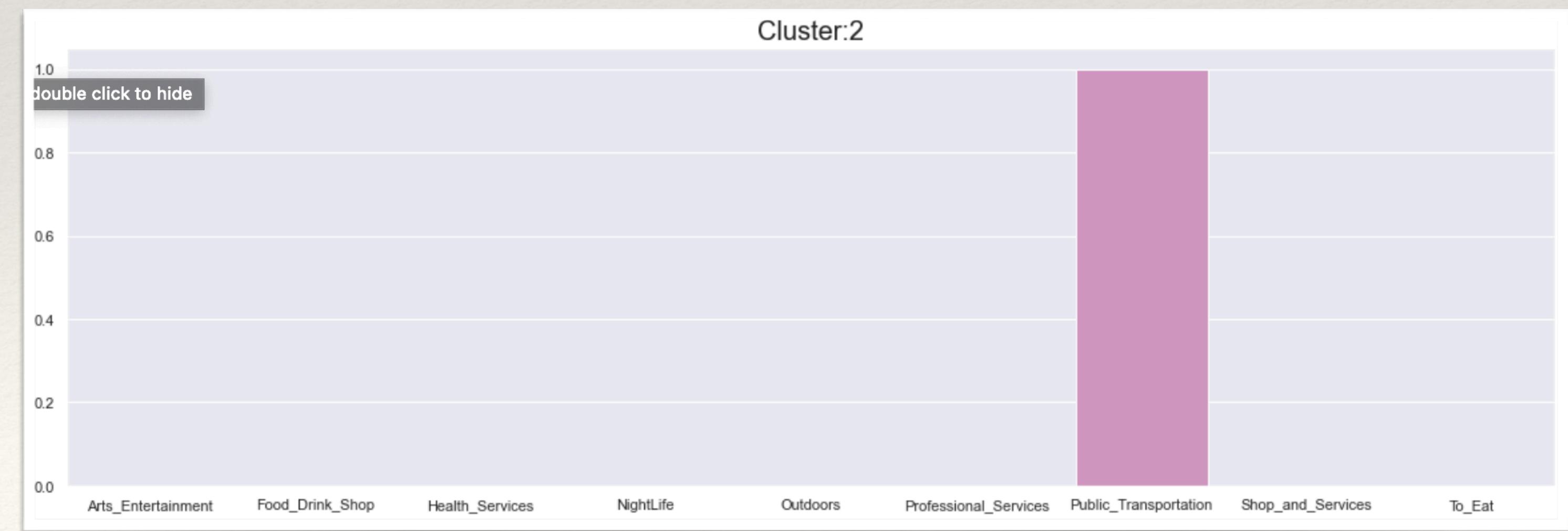
Cluster -0-



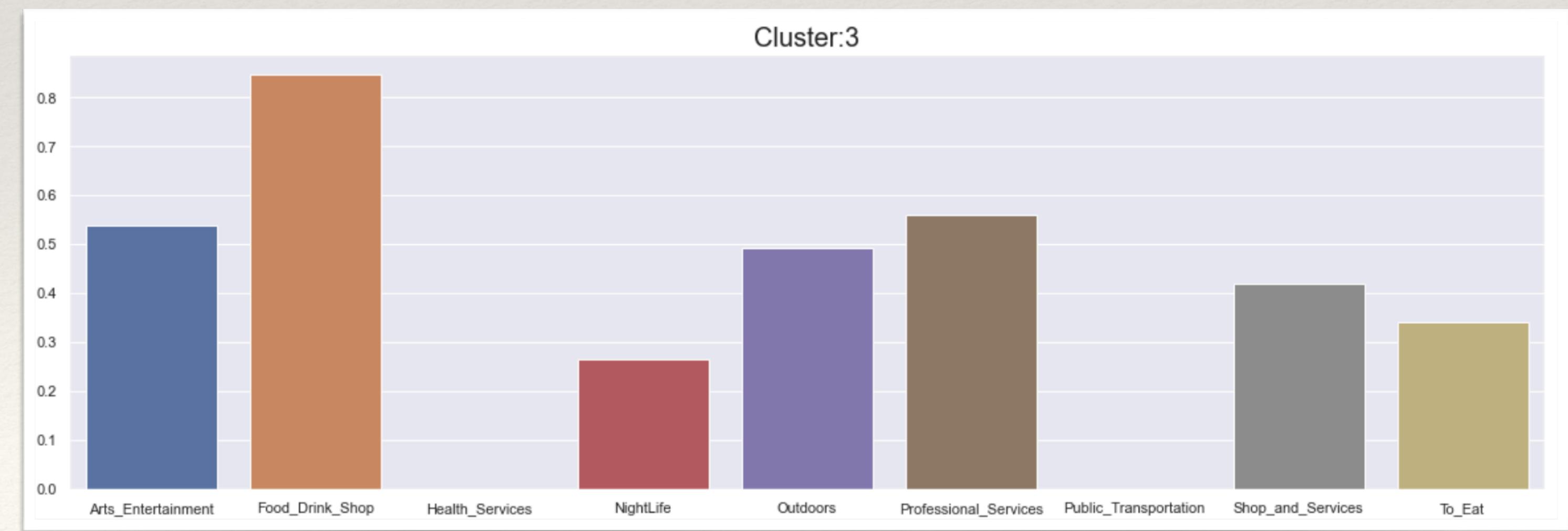
Cluster - 1 -



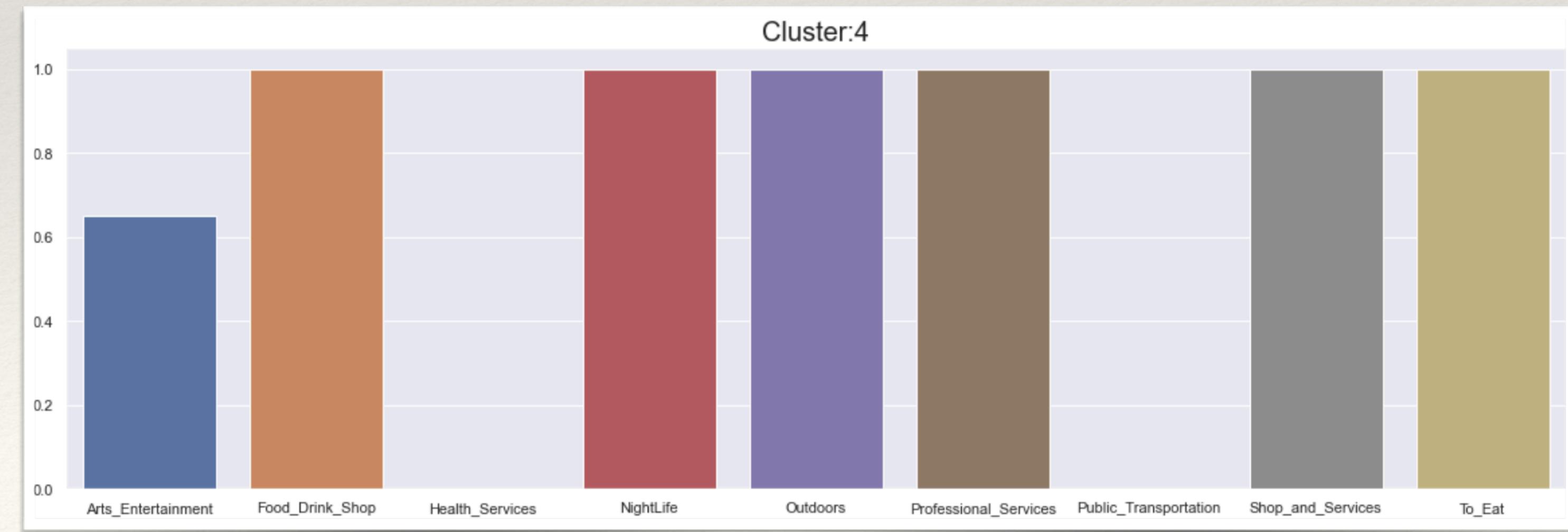
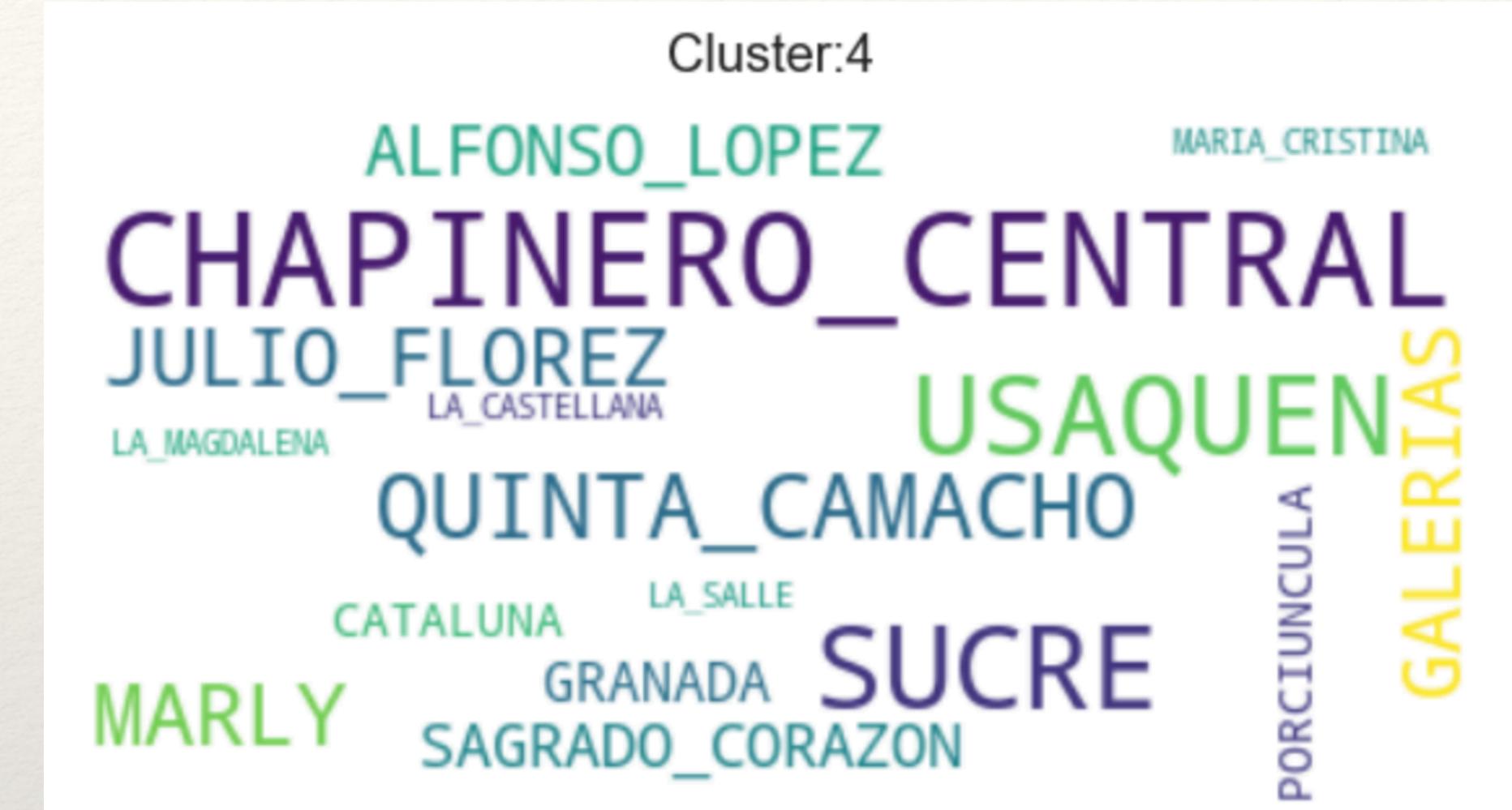
Cluster -2-



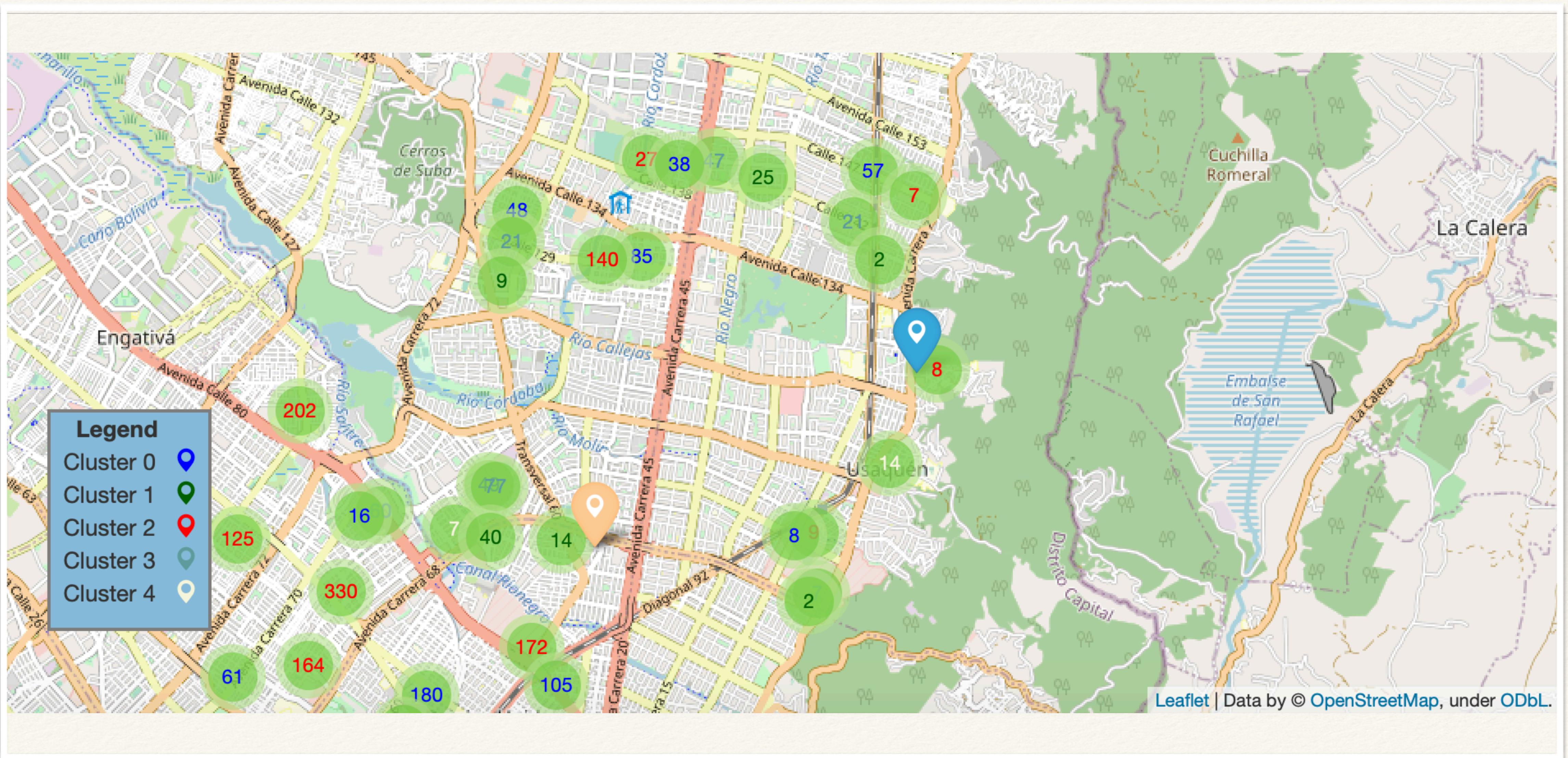
Cluster -3-



Cluster -4-



Map



Conclusions and Future Directions

- ❖ Usaquén is one of the places with the highest number of venues, which allows for a better quality of life when choosing a home.
- ❖ None of the selected points do have medical services nearby
- ❖ In this analysis, the intention was not to provide the married couple with a place to live since this implies other factors such as availability of rental housing, price, size, among others.
- ❖ For future work we could have a database of prices available for renting, this could give a greater spectrum of analysis of the areas.