

A photograph of a bakery display case. The case is filled with various breads and pastries. On the top shelf, there are several round loaves, some with sesame seeds, and some with a dark filling. The middle shelf features long, braided breads with a filling, and several round loaves. The bottom shelf shows more round loaves and some pastries. The display case is made of glass and has a black frame. The background is slightly blurred, showing a street scene with buildings and a car.

Market Research Bakery

A photograph of a bakery display case filled with various pastries and breads. The pastries are arranged on wooden shelves. The text is overlaid on the image.

Importance of the study:

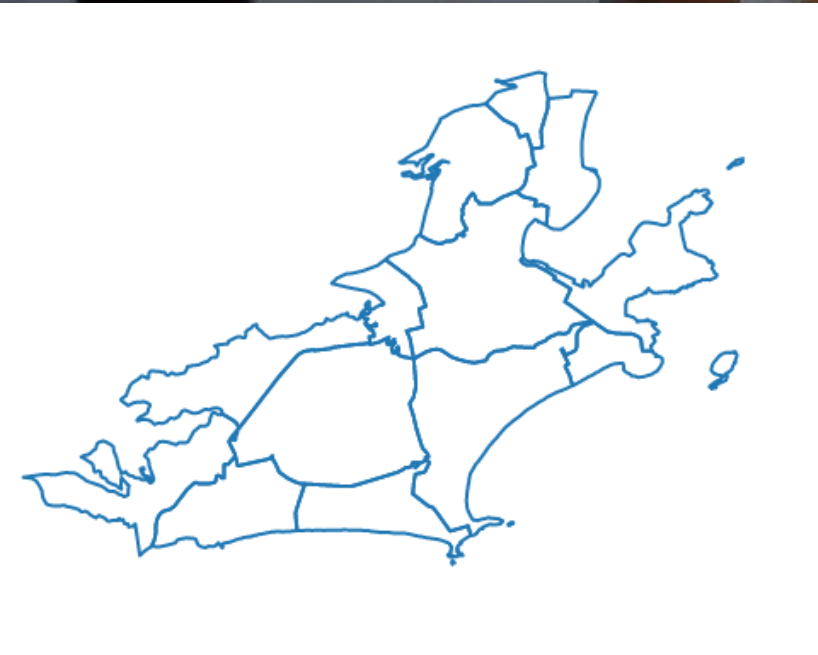
Targeted audience is mainly formed by people who live in the surroundings of the place.

Some aspects like HDI and population density have relation with the chance of a new bakery being succeed.

It's fundamental to know where are located the other bakeries in a given region so that you can avoid areas with bakeries in excess.

Neighborhoods analyzed:

The Neighborhoods of this study model will be those of the South Zone of the city of Rio de Janeiro. Neighborhoods included:



Gávea
Leblon
Ipanema
Lagoa
Flamengo
Catete
Humaitá

Laranjeiras
Jardim Botânico
Copacabana
Leme
Botafogo
Urca

Data acquisition and cleaning

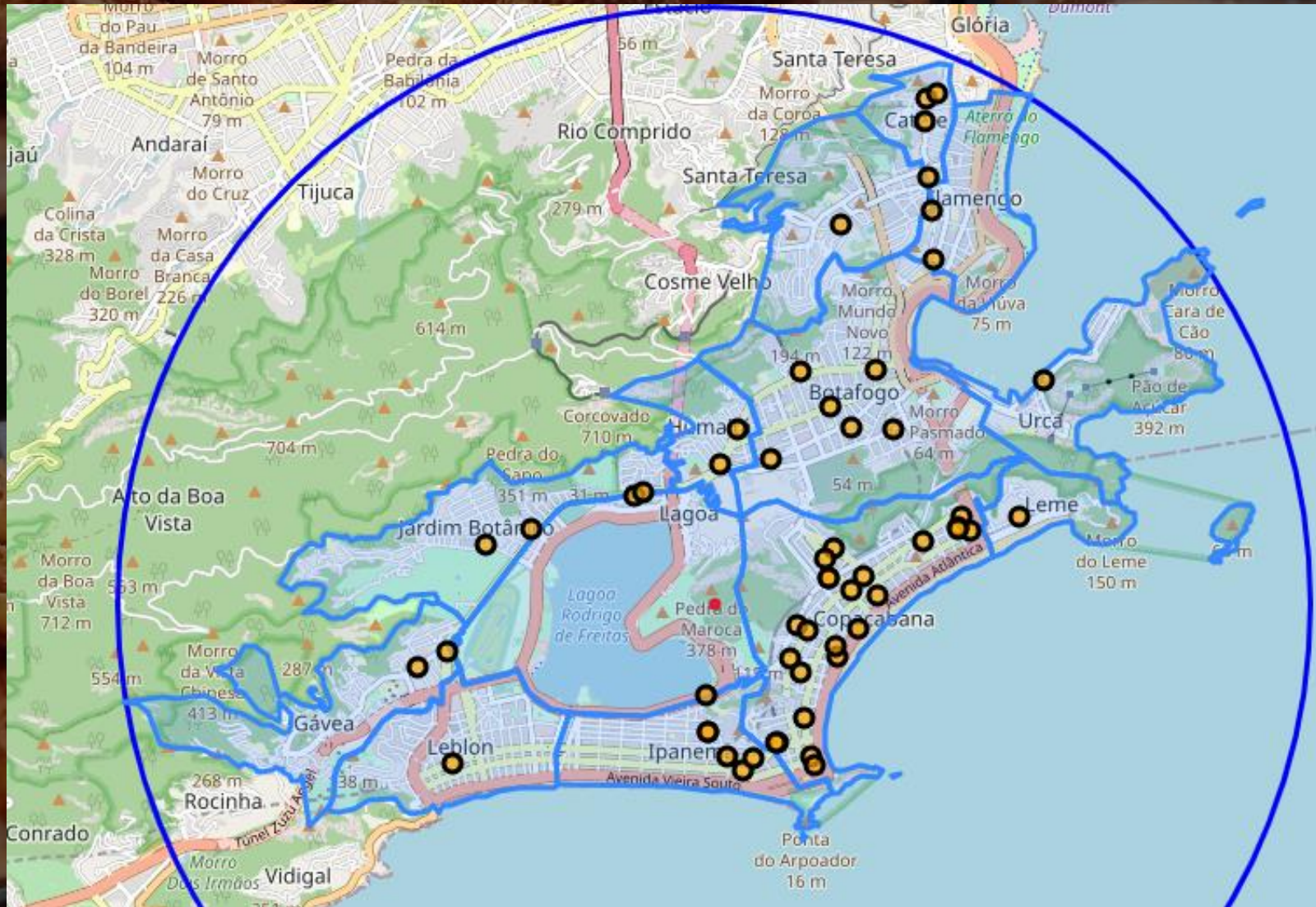
We used the API provided by the Foursquare to get information about the bakeries spread across the South Zone of the city of Rio de Janeiro

From Wikipedia, we were able to get the data of the population density and the HDI information. These data have as source the Rio de Janeiro's city hall.

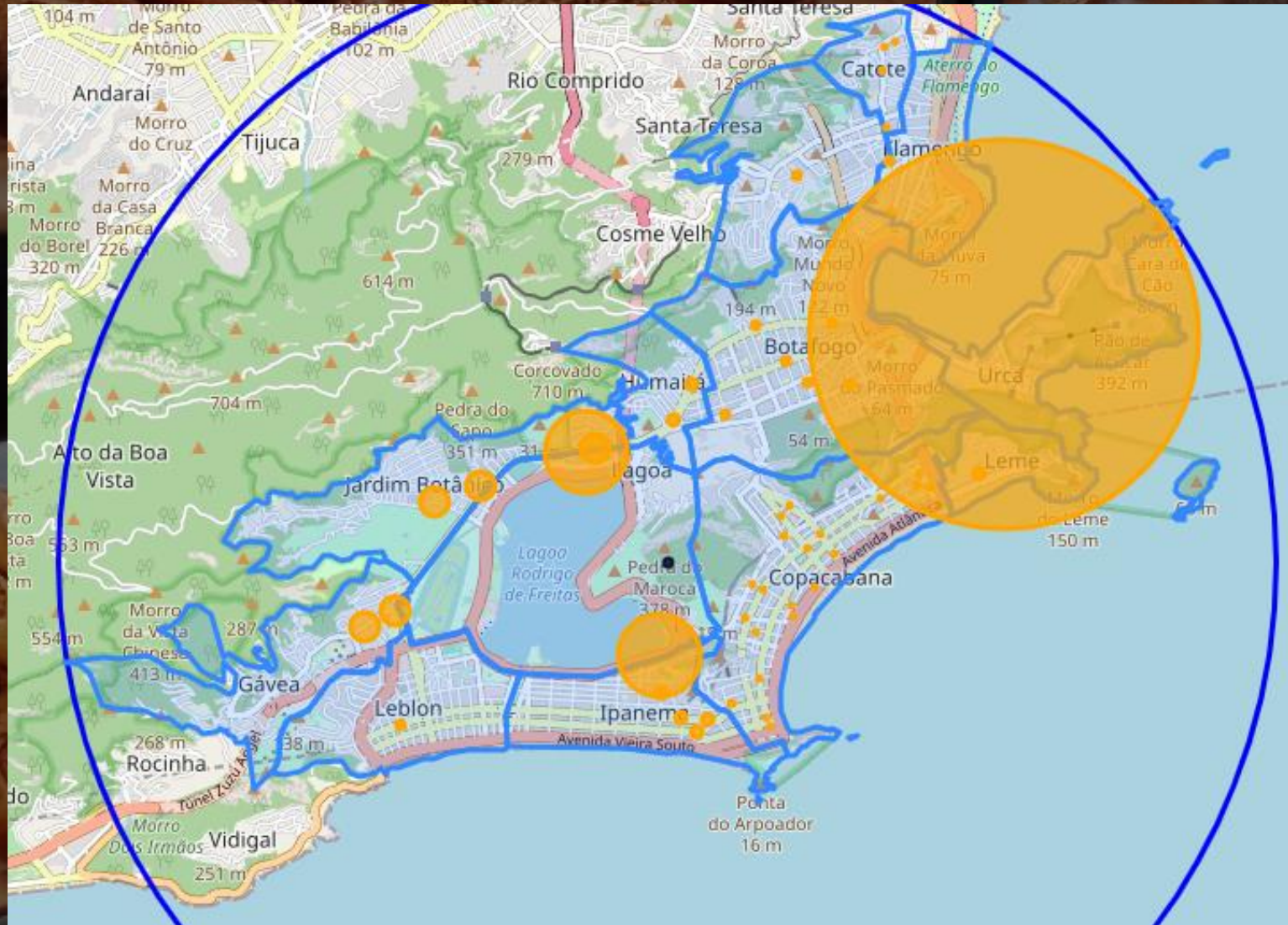


WIKIPEDIA
The Free Encyclopedia

Concentration of bakeries in the South Zone

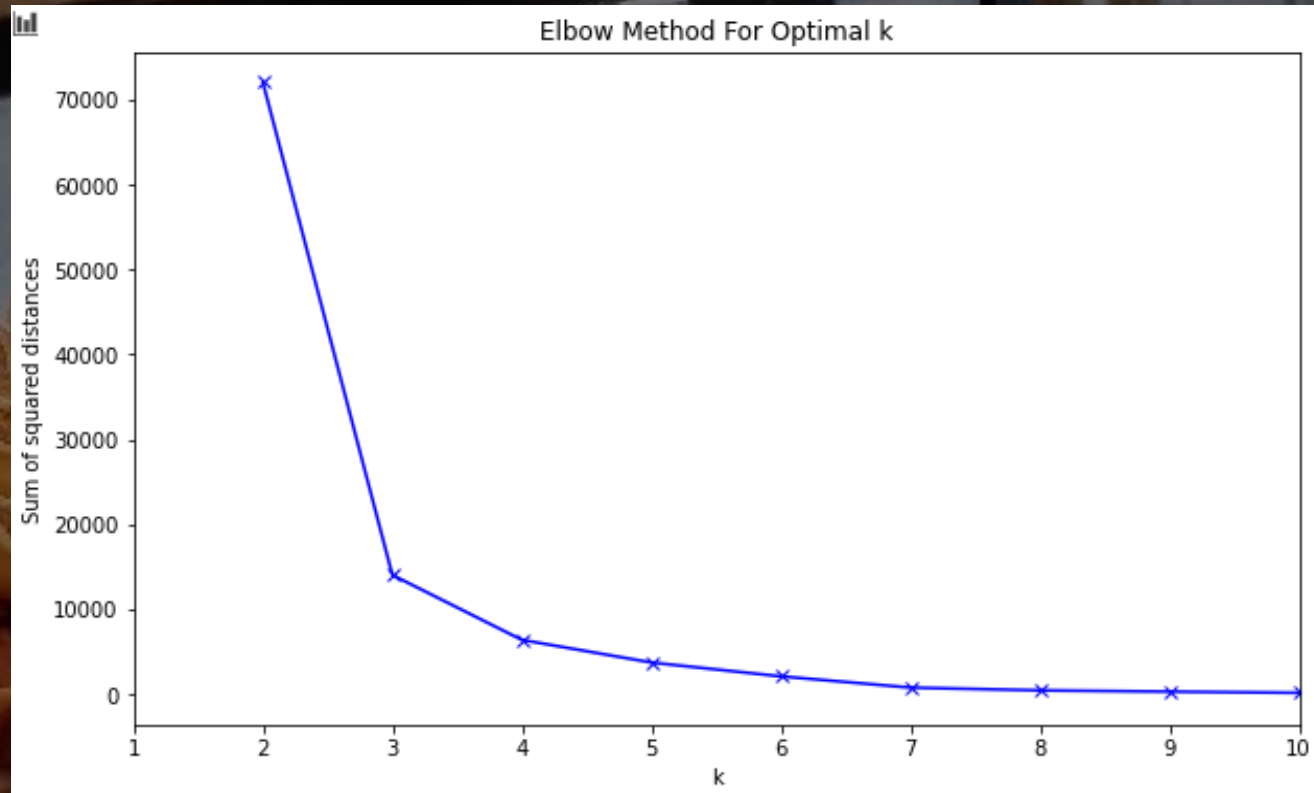


Bakeries weighted by the HDI and population density:

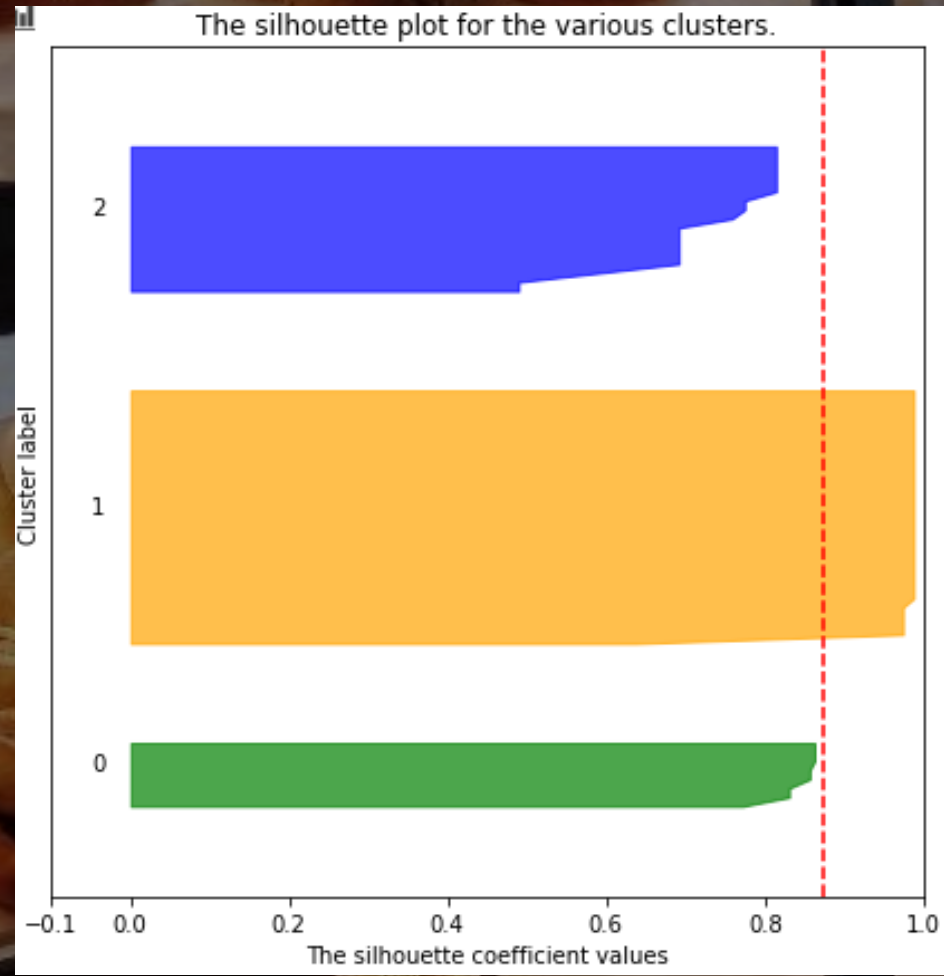


Cluster Analysis:

By analyzing the Elbow method, along with the silhouette, we can come to the conclusion that $k=3$ is the optimal value.



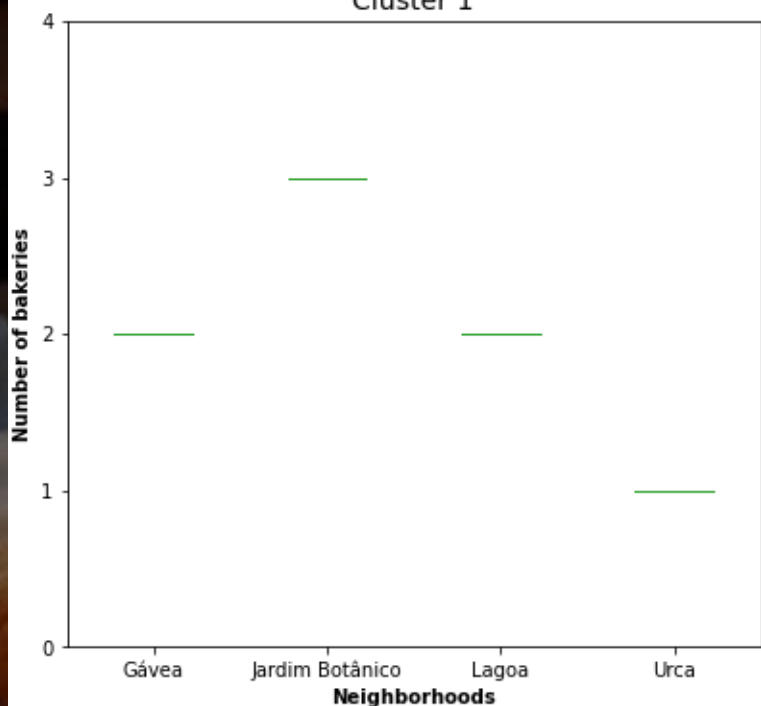
Cluster Analysis:



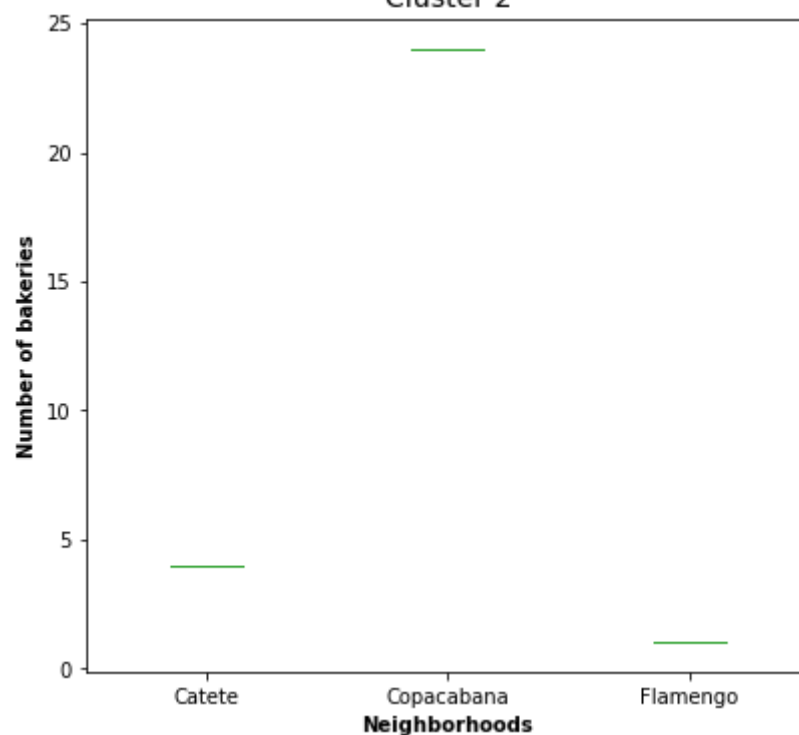
Map after applying k-means:

Comparison between clusters:

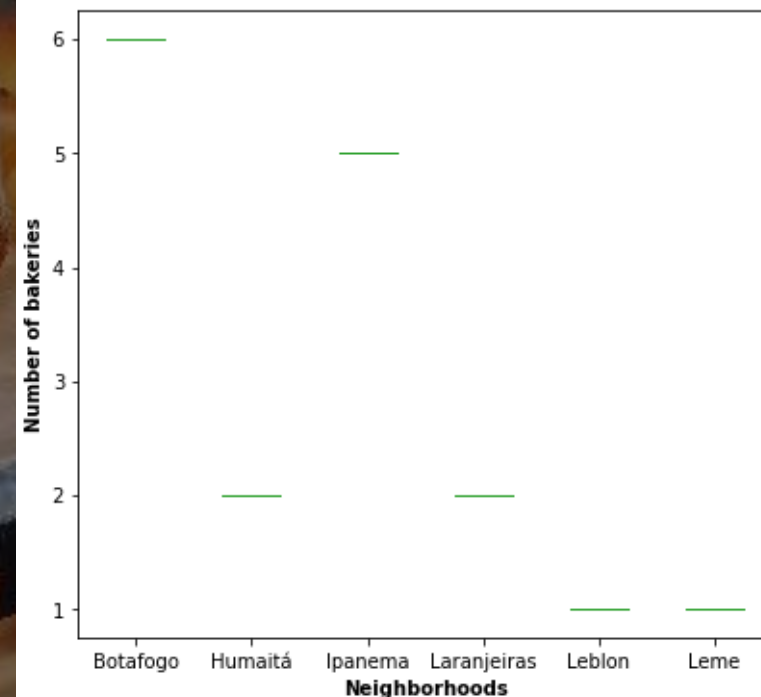
Cluster 1



Cluster 2



Cluster 3



Conclusion:

The neighborhood of Copacabana is where we have the great majority of the bakeries. Even considering that it has a greater population density and HDI, it's not a good place to start a bakery, since is the neighborhood with the highest concentration of bakeries in the South Zone of Rio.

The neighborhood of Leblon, along with Urca, is where we have the least number of bakeries, so it's a great place to start its business.



Conclusion:

By looking the clusters, we can see the following following behavior:

Cluster 1 (Green on map): It has bakeries with the lowest concentration.

Cluster 2 (Yellow on map): It has bakeries with the highest concentration.

Cluster 3 (Blue on map): It has bakeries with medium concentration.

