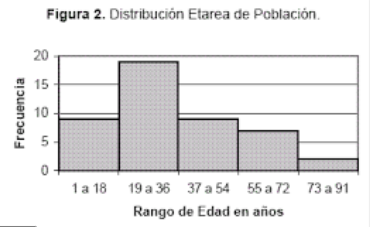
**Capstone Project: Find the best place to open a gym in Bogotá, Colombia**

**Introduction**

Bogotá is the capital of Colombia, a country with an expected GDP growth of 7% by 2021, with a young population that has become aware of the importance of healthy lifestyles in the midst of economic growth.



This is why my friend has identified a business opportunity by opening gymnasiums in strategic places in the city, which are beginning to be identified as areas where to find these types of establishments. However, it is not clear to find the best areas to start the study, therefore it is proposed to carry out a Data Science exercise to make use of machine learning and clusterig tools to deliver a solution that takes into account the configuration of the city and how it has been done. been developing.

Interfaz de usuario gráfica

Descripción generada automáticamente con confianza media

**Data selection**

To carry out this exercise, it is necessary to identify the postal codes of the different neoghborhood in the city, as well as their geographical location expressed in latitude and longitude.

Mapa

Descripción generada automáticamente

With this information, foursquare data will be used to identify the premises by area and build the grouping model.

**Methodology**

In the first place, a data cleaning is done for the table of postal collected for Bogota, where it is necessary to indicate the locality and the neighborhood to achieve a better analysis. The data is downloaded in cvs format and later uploaded to jupiterlab for manipulation and analysis.

Subsequently, the information related to geographic location with latitude and longitude is reviewed and a map is plotted to verify the consistency of the information.

Then, a foursquare download is made with the premises by postal code classified according to the corresponding latitude and longitude, asking for the 10 most representative premises of each neighborhood and town.

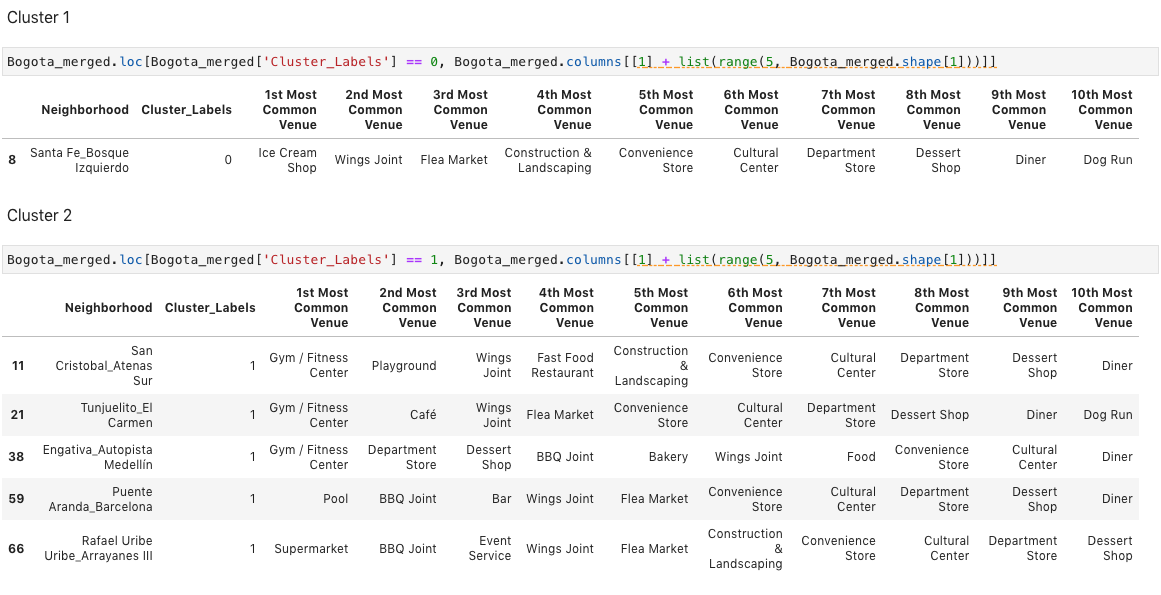
Finally, it is built with a clustering model to identify the best areas where the problem raised can be solved.

With the result of the model, the areas where potentially a better commercial performance would be achieved for the location of the gymnasiums are identified.

**Results**

Mapa

Descripción generada automáticamente



Tabla

Descripción generada automáticamente

Interfaz de usuario gráfica, Texto

Descripción generada automáticamente

**Discussion**

The city has 5 clusters with different characteristics, where it was possible to detect one where gymnasiums and fitness places that are potentially suitable for entrepreneurship predomínate

**Conclusión**

In conclusion, cluster 2 presents localities that are highly oriented to health establishments, gyms and fitness, therefore it is proposed as an appropriate area to start the search for places where new gyms can be opened since, due to its configuration and characteristics, it already attracts to the target public of this endeavor.