SPORT NUTRITION STORE @ GRANADA, SPAIN

IBM DATA SCIENCE COURSERA CAPSTONE FINAL PROJECT

Introduction

The objective of this capstone project is to analyze and select the best location in the city of Granada, Spain to open a new **sports nutrition store**. Using data science methodology and machine learning techniques like clustering, this project aims to find what is the best location in Granada to start a sports nutrition store and to persuade investors to fund this business.

Data

The main sources of data are the following:

- List of neighborhoods and municipalities in Granada

- Latitude and longitude coordinates of those locations

- Venue data, particularly data related to Gyms

Methodology

▶ Web scraping Wikipedia page for neighborhoods list

▶ Get latitude and longitude coordinates using Geocoder

▶ Use Foursquare API to get venue data

▶ Group data by neighborhood and taking the mean of the frequency of occurrence of each venue category

▶ Filter venue category.

▶ Perform clustering on the data by using k-means clustering.

▶ Visualize the clusters in a map using Folium.

Results
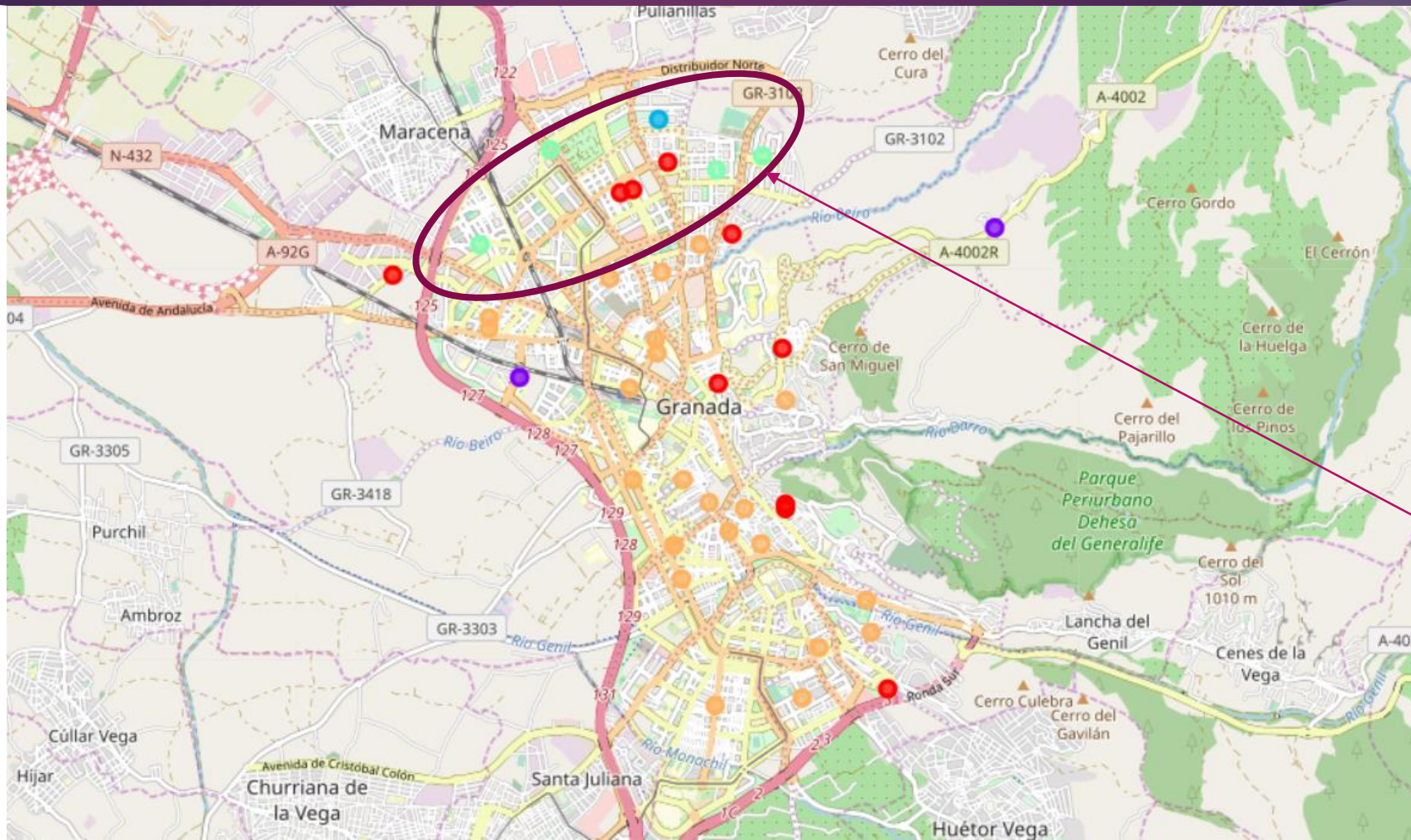
CLUSTER OF INTEREST

Discussion

After see the results, we can highlight some points:

- Most of the gyms are concentrated in cluster 3.

- This represents a great opportunity and a high potential area to open a sport nutrition store as there is a big number of potential clients, due to the density of gyms in the area.

- There could be also opportunities to open a gym the rest of neighborhoods.

Conclusion

The neighborhoods in cluster 3 are the most preferred locations to open a new sport nutrition store as the have the higher number of gyms. The findings of this project will help the relevant stakeholders to capitalize on the opportunities on high potential locations while avoiding overcrowded areas in their decisions to open a new sport nutrition store.