1. INTRODUCTION – BUSINESS PROBLEM

The objective of this project is to help investors to find the best area for opening a restaurant in Madrid, Spain.

Madrid is the largest Spanish city with more than 3 million inhabitants in a metropolitan area of approximately 600 square kilometers, and it is indeed a very good place to open a restaurant, thanks to its busy business centers, and lively streets either by day or night.

Madrid is administratively divided in 21 districts ("distritos" in Spanish), and 129 wards ("barrios" in Spanish).

Thanks to the data from several Madrid City Hall webpages and Foursquare, and an appropriate data science approach, at the end of the project stakeholders will have the proper information to select the best wards and the best kind of restaurant to invest in.

2. DATA

Based on the objective of the business problem, the following data will be retrieved:

- Information of the administrative distribution in district and wards. This information will be obtained from the Madrid City Hall official website
- Information of all of the district such as population, size (area), average income, number of crimes, or number of hotel beds.
- Information of the coordinates of every ward by geocoding
- Information about the type and location of restaurants from Foursquare API

3. METHODOLOGY

In order to accomplish this objective, three main phases will be considered:

1. Selecting the best "distritos" to open a restaurant. This selection will be done by clustering the "distritos based on its size, population, average income, number of crimes per year, and number of hotel beds (as a good indicator of tourism).

2. Selecting the best "barrios" to open a restaurant within the best "distritos" selected in phase 1. This selection will be done by clustering the "barrios" based on the number of restaurants and type.

3. Selecting the best three types of restaurant to invest in, considering the clusters created in phase 2, based on the less common kinds of restaurants of each cluster.

4. RESULTS

5. DISCUSSION

6. CONCLUSION

7. APPENDIXES