

Project Description DS4A - Team 15

Project 38: Procolombia

Overview of ProColombia

Advertising exports of non-mining energy goods and services in potential markets, expansion of Colombian companies, attraction of foreign investments, positioning the country as a tourist destination for vacations and meetings, and Colombia's "Country Brand", are the main axes on which PROCOLOMBIA focuses its work¹.

Problem Description

For ProColombia it is necessary to identify companies with potential investment in the country. Therefore, every year a segmentation exercise is carried out by its commercial area to identify new enterprises that might be interested in the country.

At the moment, ProColombia is using multiple sources of information such as *Ministerio de Tecnologías de la Información y las Comunicaciones*, information provided directly by the companies, and additional resources listed by investors. However, this exercise requires too many resources (man-hours) compared to the final results because it is necessary to match the different databases to guarantee that the companies' consolidated information is accurate.

and consumes too many resources. Also, the entity does not have the operational capacity to incorporate a statistical model in this exercise, to highlight companies that are more likely to invest in Colombia.

las cifras de empresas que nos proveen, otras empresas que listan las inversionistas. El desarrollo del ejercicio consume demasiados recursos en horas hombres en relación con el producto entregable porque emparejar o hacer match entre las diferentes bases de datos requiere de algoritmos que garanticen que se está uniendo la información de la misma empresa.

Currently Procolombia does not have the operational capacity for this analysis, so it is done with simple tools such as MS Excel. Therefore, the use Python which is more robust and flexible.

Once the different source information are joined, we expect that the system runs a statistical or econometric model in order to identify the companies that they are more likely to invest in Colombia.

The companies in the list are a strategic product for us, it allows us to know the international companies with investment potential in Colombia or potential buyers of different Colombians products.

¹ source:<https://procolombia.co/>

Why does this problem matter?

It is important because Colombia requires to attract foreign investments in order to get resources to transform its productive and economic environment. Identifying high potential companies would be very valuable for the country and support the national economy reactivation via employment, poverty reduction and improve local production. Taking this into account, it is important to optimize the manual processes that are being executed related to data extraction, storage and product development. In this way, Procolombia will have more time available and focus on obtaining insights to capture the best opportunities, and have a positive impact in Colombia.

Datasets

The following datasets are provided by ProColombia:

1. Database of international investment companies : Company name, Industry, Financial variables (Income, expenses, profit, etc.), locations, other variables
2. Database of companies that export to Colombia: Company name, industry, Exports frequency and values, other variables
3. Database of companies established in Colombia. :Company name, industry, total investments, number of employees, other variables
4. Database of companies with investment signals around the world: Company name, industry, total investments, other variables.

Additional databases (external datasets)

5. Database of historic investments (exports and imports)
6. Database of International Financial Reporting Standards.