

Executive Summary

Challenge

Track wine prices on online stores to guide Sogrape's distribution and pricing strategies. For that, we used Web Scraping to harvest the data that was treated and stored at a database to provide useful information for a crating of a visual and clear dashboard.

Difficulties and solutions

- At first, the problem was about the Web Scraping approach on how it would collect the website information in a fast, legal and ethical way. The first try ended up to be an overload of information, so we redesigned the Web scraper to collect the information of the specific EAN on the listed stores.
 - Garrafeira Soares website redirected the Web scraper into a page that only had a script, so it had to be coded to clean the information and find the product page.
 - El Corte Ingles Supermarket has a protection against scraping, so the Web scraper fakes an user's browser to overcome the protection.
-

Strengths

- The Web scraper is coded to find the information about a specific product on the listed websites one at a time, sequentially, creating an artificial delay to prevent multiple requests on a single website at once.
 - The dashboard allows the user to simply add a product to the scraper to harvest it's information only by filling up the name of the wine and the EAN.
 - Can easily be configured in a cronjob to be executed at the chosen date and time, and the existence of a script to automatically include the required dependencies.
-

Roadmap

- Adding a profile dashboard completely personalized to the user's role in the organization that contains objective information about the products of interest. Include price alerts to each brand manager.
- Include graphics that compare product's price on the specific stores.
- Improve the dashboard filters with information about other countries and different postal codes.
- Scrape and display compared information about competitors.
- Add documentation for the implementation and maintenance of the project.
- Implement a docker to easily deploy the application in other locations.