

Challenge / Opportunity

Between the afternoon of the 23rd and the morning of the 26th October 2023, Sogrape and 42 Porto hosted a Hackathon under the following theme: “Online Wine Price Harvesting Challenge”.

The challenge the participants were tasked with was implementing a web scraping system able to collect wine prices from national online stores on a daily basis, and to develop an interactive and intuitive interface to display the same data.

Approach to the challenge

We are presenting an automated web scraping website that will allow Sogrape’s employees to easily compare pricing between several retailers for the same product, and to keep track of price fluctuations over time. Our website effectively addresses compatibility and portability issues, not requiring any installation.

What we currently offer

- We are handling 3 national retailers, plus an international one (Spanish branch of El Corte Inglés), knowing that we are capable of expanding our web scraping tool in the near future.
- In order to effectively extract data from websites that are built differently, we are proposing a solution that addresses each of the predefined target websites in an adequate way, that is, being able to deal with the different classes used in the way each one of them is structured.
- Using SQL to support our database proves to be a scalable solution, future-proofing our product.
- We have designed a clean-looking easy to use interface to filter and see collected data, providing valuable insights and making way for better informed data-driven decisions

Room to grow & improve

- Add more websites to scrape from
- Filter data with more than one filter at once through the graphic interface
- Include more filtering options
- Add more visualization options (e.g. price fluctuation history graph), making data tracking more immediate
- Allow the user to input data (e.g. product name) through text in a search bar, that will be useful when navigating bigger amounts of data