



Hackathon • Team 10

S O G R A P E



42

| PORTO

Online Wine Price Harvesting Challenge

# The Team

Hackathon · Team 10



**Bernardo Esteves**  
bbento-e



**Bruno Lopes**  
brpereir



**João Ramalhosa**  
joaoped2



**Ricardo Santos**  
ride-sou



**Teresa Chow**  
tchow-so



**Vinicius Hehn**  
vivaccar



# Approach

1. The Challenge
2. Strategy
3. Technical dimension
  - Web Scraping
  - Data Management
4. Visualization & Interaction
  - Interface Design
5. Current state
6. Next steps





# The Challenge

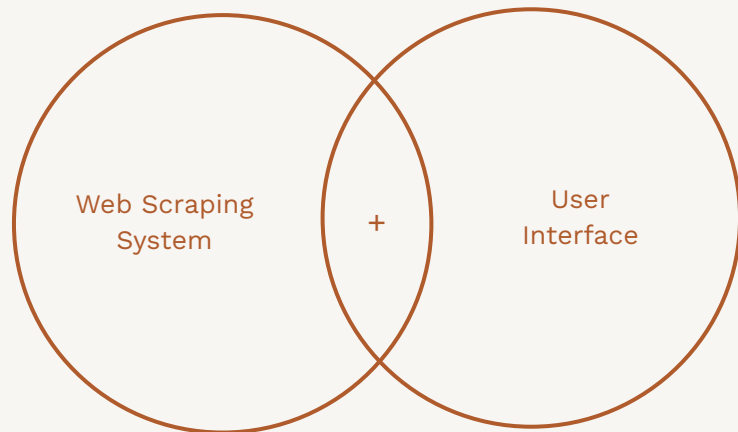


# The Challenge

Design and implement a web scraping system and respective interface to keep track of product prices practiced across various stores and markets.

In its 1st stage, it should:

- collect data from **3 websites...**
- ...relating to **4 products**
- present collected data in a **user-friendly dashboard**



# Strategy

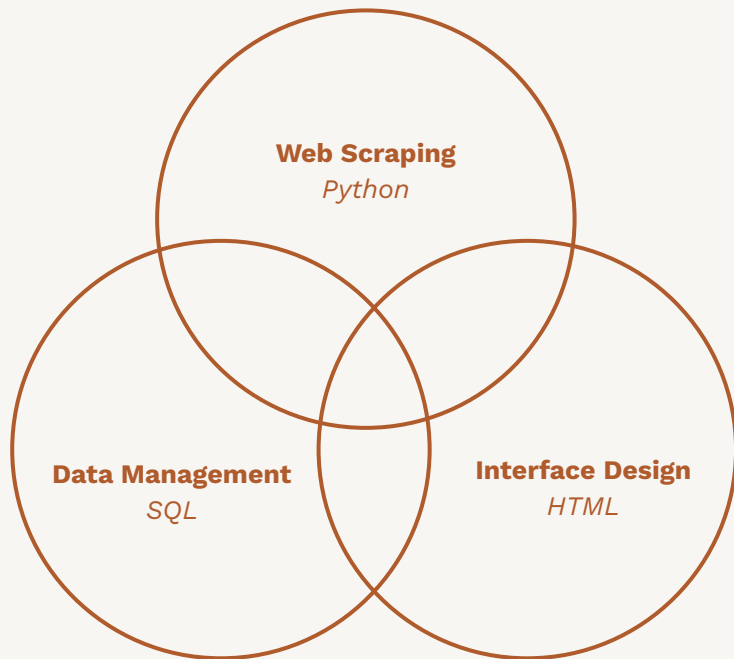
A grayscale photograph of a person with curly hair and glasses, seen from the side, sitting at a desk and working on a computer. The desk has two large monitors. The left monitor displays a code editor with a dark theme, showing a file explorer on the left and code in the main area. The right monitor is partially visible. The person is wearing a dark shirt and is looking at the screens. The background shows a window with a view of a city building. The word "Strategy" is overlaid in a bold, orange font in the center of the image.

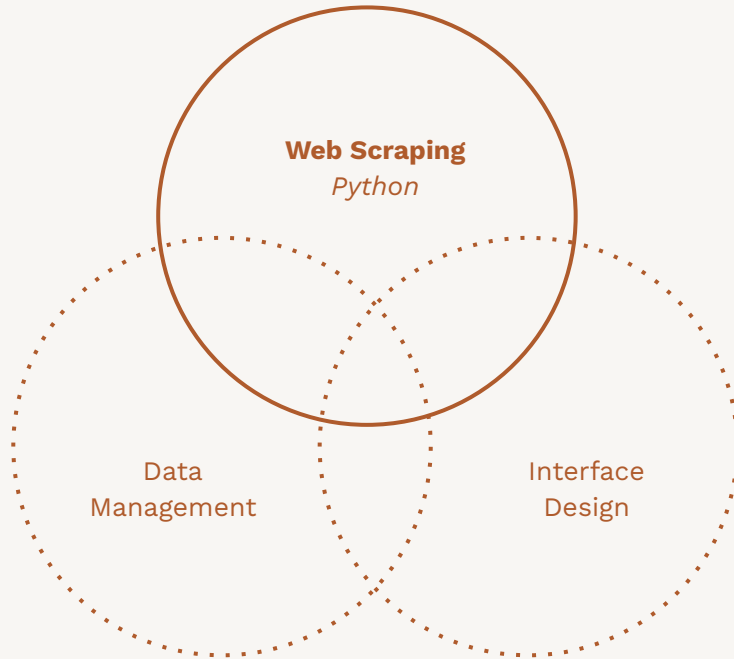


# Allocating resources

We decided to subdivide the Team into three action fronts,

- **subdividing** a bigger problem into smaller and **more manageable** ones that could be **solved independently** and **simultaneously**
- **combining smaller solutions** in order to tackle the original challenge
- ensuring **constant communication** within the team so that **everything fits**





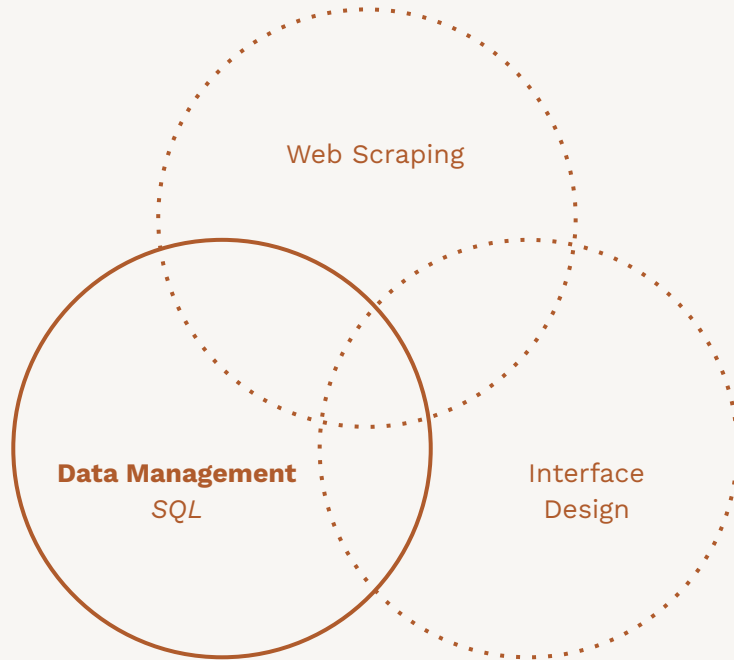
- **data analysis**

data → valuable insights

- identify target website(s)
- use locators to find relevant information in the HTML,  
different websites = different classes

extracted data → delivered in structured format

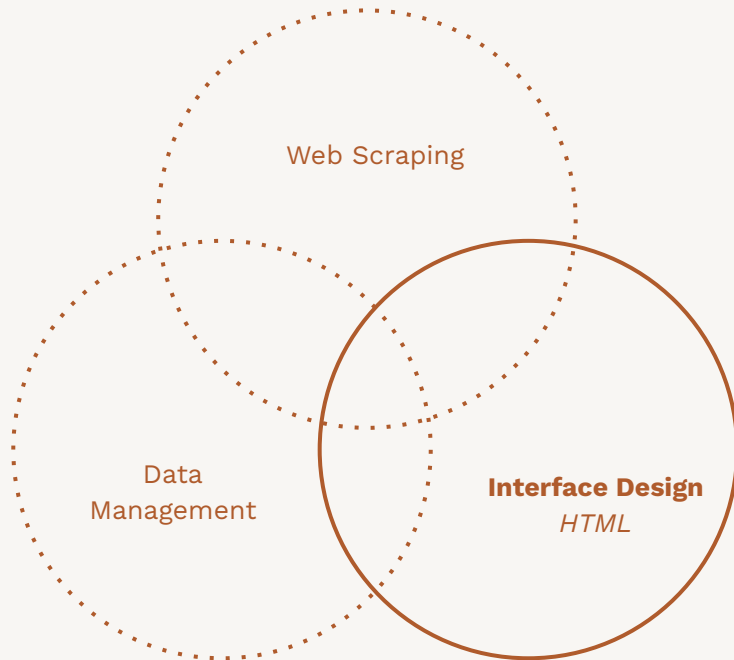




- **data storage, retrieval, management & manipulation**

SQL ⇔ relational databases

- standard & reliable
- easily scalable, allowing room for future improvement



- **user interface**

user input  $\rightleftharpoons$  data visualization

support the effective functionality of website:

- facilitate interaction between (non-tech) user and program
- presenting data in a visual manner
- filter data



# Current state

What we are currently delivering:

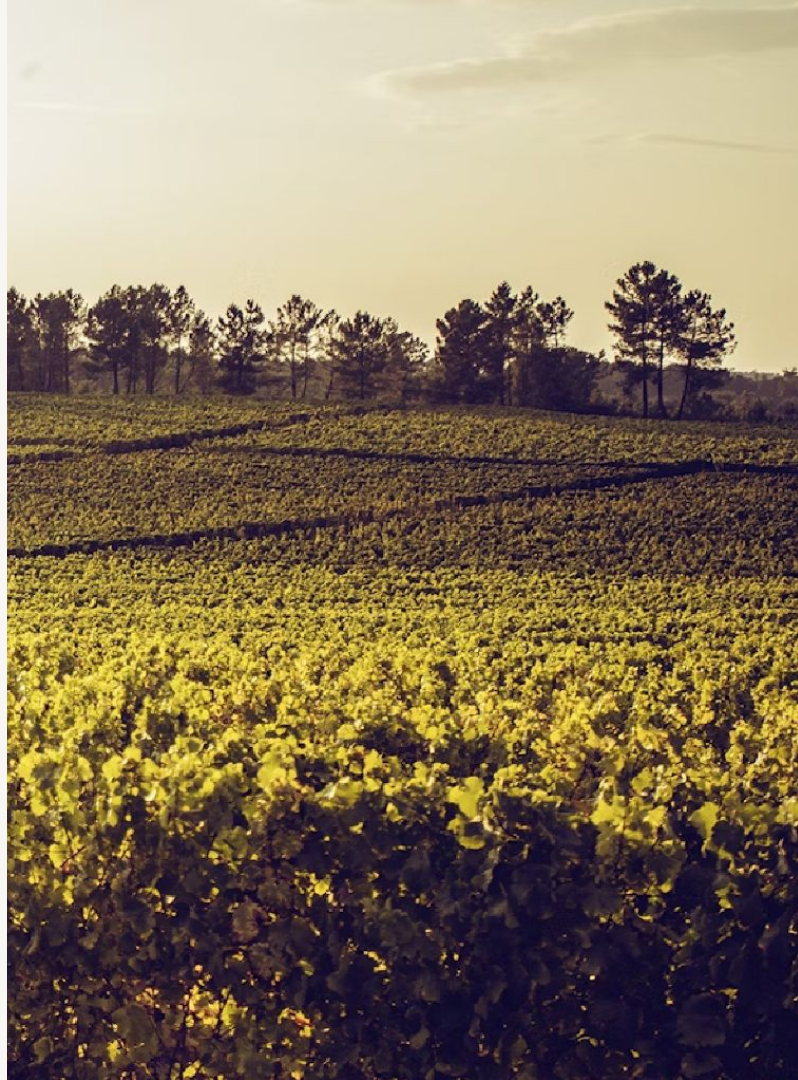
- automated tool, capable of filtering:  
product's name, retailer and location
- valuable & actionable data insights
- portable & compatible  
(website vs. app)
- intuitive user interface  
(accessible to non-tech profiles)





## Next Steps

- 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







Thank you