# Technical Test Deporvillage - Data Scientist

You are required to generate a prediction of the total amount to sell during the year 2024-09-01 → 2025-08-31 for each brand+family. The goal of this prediction is to generate a budget (per brand+family) for the purchasing team.

You are given two data files, one with information about products, and another containing historic sales data.

Please hand in one (or several, if required) `ipynb` file with the data exploration, modelling and outcome analysis; and a data file with the generated predictions.

The developed ML solution does not need to be finetuned, since the goal is to evaluate your approach to the process, not the specific result.

Then answer the following questions:

1. What metric have you used and why have you selected this particular metric? Is there any drawback that comes from using this metric?
2. What error can we expect from the generated predictions? How many units do you estimate will be left unsold at the end of the year? How can we minimize this?
3. If you had the time to develop this project further, what improvements would you consider testing?
4. How would you make these predictions available to the purchasing team? Please write a few pros and cons of every alternative proposed.
5. How would you evaluate the performance of this model in production? How would you justify those numbers to the purchasing team?
6. If there is any pandemic or economical crisis, which strategy would you propose to make sure the model can be adapted to it?