Data Scientist - Technical Test

We are working on a lead scoring algorithm for a client that has an Event Management SaaS application. Your job is to receive the raw data from the client and prepare it to train a classification algorithm that will predict the probability of conversion of each lead.

You will receive 2 datasets that reflect the phases in the client's sales process: **leads.csv** (data of all potential clients) and **offers.csv** (clients that got at least to the demo meeting). The target column will be Status from the offers dataset, which indicates whether a client has bought the product.

Deliverables:

1. A single enriched dataset that will guarantee the best results when we train the classification algorithm (you are expected to do feature engineering)
2. The Python script used to process the data
3. Some considerations on the data and the problem, and a brief explanation of your thinking process presented as a simple Data App created with our SDK. See here:
    - https://github.com/shimoku-tech/shimoku-api-python
    - https://docs.shimoku.com/development/getting-started/quickstart