

faire.ai's Data scientist case study

Hi and welcome to the faire.ai's case study that we hope you'll enjoy to develop with your best skills and creativity.

1. Create a **Python Notebook** on [Google Colaboratory](#) that:

a. Create a Synthetic Dataset having 5k points:

Variable	Distribution
Sex	$Bernoulli(p = 0.4)$
Age	$Gamma(\alpha = 8, \theta = 4)$
Purchase	$Unif(\alpha = 500, \beta = 550) - 10 * Age * Unif(\alpha = -2, \beta = 0.01)$
Output	$Bernoulli(p = 0.3 * Sex + (0.2 \text{ if } Age > 30 \text{ else } 0.4) + (0.3 \text{ if } Purchase > 100 \text{ else } 0))$

b. Create a Model to predict the **Output** variable

2. Using your Data Science experience explains the decisions taken in the project.

How will the project be evaluated?

- Clear code
- Deep documentation
- Apply correctly concepts from Statistics and ML
- Delivery Time