

## Descrição do case

O time de *analytics* do C&A Pay está planejando uma redução na aprovação de clientes para reduzir custos com perdas. Para isso, o time de modelagem pensou em utilizar o histórico de compra do cliente na C&A para tentar criar um modelo de *application* que consiga diferenciar clientes propensos à perda na concessão de crédito ao cliente.

Dados: Os dados para análise foram quebrados em três tabelas. Abaixo segue descrição das respectivas tabelas e campos:

1. `case_target.csv`: esse arquivo contém uma amostra de clientes que fez o c&a pay em determinadas safras com a posterior avaliação de inadimplência.
  - a. `id_cliente`: chave primária do cliente
  - b. `target`: 1 para inadimplente e 0 para adimplente.
2. `case_compras.csv`: esse arquivo contém dados de compras dos clientes da amostra na C&A posteriores à concessão do C&A Pay.
  - a. `id_venda`: chave primária de venda
  - b. `id_semana`: Id da semana
  - c. `id_cliente`: chave primária do cliente
  - d. `id_loja`: Id da loja
  - e. `id_item`: Id do item
  - f. `id_item_cluster`: Id do cluster do item
  - g. `id_item_departamento`: Id do departamento do item
  - h. `id_item_divisao`: Id da divisão do item
  - i. `id_item_familia`: Id da família do item
  - j. `vl_item`: Valor do item
  - k. `qt_item`: Quantidade do item
3. `case_pgto.csv`: esse arquivo contém dados de pagamento das respectivas compras dos clientes da amostra.
  - a. `id_venda`: chave primária de venda
  - b. `id_meio_pgto`: Id do meio de pagamento
  - c. `vl_pgto`: Valor do pgto

O desafio deve ser entregue com o código, podendo ser em R ou python. Fique à vontade para avaliar as variáveis e desenvolver o modelo com qualquer método que achar fazer sentido.

Caso fique com alguma dúvida, entre em contato via linkedin:

<https://www.linkedin.com/in/caio-paes-carneiro-934a529b/>