

## **Desafio Técnico 1 - (System Design)**

A ideia deste exercício é entender a sua visão de arquitetura de dados que possibilitem possibilitam resolver esse problema de negócio:

Dadas APIs não sincronizadas não sincronizadas que enviam eventos, ou seja, APIs que enviarão eventos em timestamps diferentes sobre informações dos nossos varejistas (cupons, ou seja, vendas, carga de novos operadores, carga de novos produtos, etc).

O objetivo do sistema que deve ser construído é ingerir e processar esses dados como uma série temporal executando um modelo de ML que descobre quais são os possíveis produtos que alguém vai comprar no futuro (entrega de valor para negócio).

O processamento deve ser executado diariamente com os dados do dia anterior e o resultado deve estar pronto o mais cedo possível (performance).

Com as informações de produtos, o sistema precisa ser capaz de usar esses resultados para enviar campanhas para todos os consumidores da base com o melhor produto através do melhor canal (whatsapp, SMS, e-mail ou Push Notification), pense aqui em como calibrar os canais.

A solução precisa descrever qual arquitetura poderia resolver essa necessidade, incluindo o treinamento do modelo, componentes da arquitetura e o versionamento.

Dicas:

- Explique por que as ferramentas são as mais adequadas para a solução.
- Pense e descreva os detalhes da solução, não apenas a visão geral da arquitetura.

## **Desafio Técnico 2 - (Code Design)**

A ideia desse exercício é entender sua habilidade com codificação e entender como você resolveria um problema como esse que será seu dia a dia como Engenheiro de Dados dentro da Mercafacil. e nada melhor do que começar a entender um pouco mais do mundo mercadológico. Confeccionamos esse desafio prático para que você com base nos datasets a seguir, construa os seguintes modelos:

- 1. Lista dos 5 melhores produtos que o cliente nunca comprou (modelo de Up-sell de produtos)**
- 2. Lista dos 5 melhores produtos que mais são vendidos com outro produto (modelo de Cross-sell)**

A conclusão de todos os modelos não é um requisito obrigatório.

No dia da nossa reunião de code-review compartilhe conosco a sua lógica e como fez para alcançar esse objetivo

Link para download dos datasets:

[https://drive.google.com/drive/folders/1z0cve4RKTzJGGAb4gl8-F5d2\\_ea-iZGO?usp=share\\_link](https://drive.google.com/drive/folders/1z0cve4RKTzJGGAb4gl8-F5d2_ea-iZGO?usp=share_link)

**Descrição dos datasets:**

Vendas:

CAMPO	DESCRIÇÃO
COD_ID_LOJA	Código identificador da Loja
NUM_ANOMESDIA	Data de compra no formato yyyyymmdd
COD_ID_CLIENTE	Código identificador do cliente
DES_TIPO_CLIENTE	Código identificador tipo de cliente (F para pessoa física ou J para jurídica )
DES_SEXO_CLIENTE	Código identificador sexo do cliente (M ou F)
COD_ID_VENDA_UNICO	Código identificador da venda
VAL_VALOR_SEM_DESC	Valor sem considerar desconto
VAL_VALOR_DESCONTO	Valor do desconto
VAL_VALOR_COM_DESC	Valor considerando desconto
VAL_QUANTIDADE_KG	Quantidade
COD_ID_PRODUTO	Código identificador do produto

Clientes:

CAMPO	DESCRIÇÃO
COD_ID_CLIENTE	Código identificador do cliente
DES_TIPO_CLIENTE	Tipo cliente
NOM_NOME	Nome do cliente
DES_SEXO_CLIENTE	Código identificador sexo do cliente (M ou F)
DAT_DATA_NASCIMENTO	Data de nascimento

Produtos:

CAMPO	DESCRIÇÃO
COD_ID_PRODUTO	Código identificador do produto
COD_ID_CATEGORIA_PRODUTO	Código identificador da categoria
ARR_CATEGORIAS_PRODUTO	Array com todas as categorias relacionadas de pai para filha
DES_PRODUTO	Descrição do produto
DES_UNIDADE	Unidade de medida
COD_CODIGO_BARRAS	código de barras

Categorias Produtos:

CAMPO	DESCRIÇÃO
COD_ID_CATEGORIA_PRODUTO	Código identificador da categoria
DES_CATEGORIA	Descrição da categoria
COD_ID_CATEGORIA_PAIS	Código identificador da categoria pai