

# Data Science iFood - Case de avaliação Técnica

## Visão Geral

Para aumentar o engajamento dos clientes, as empresas fazem uso de cupons que são enviados por meio de vários canais de marketing. Essa estratégia não apenas atrai os clientes para um serviço, mas também pode estabelecer um relacionamento de longo prazo, aumentando a probabilidade de transformar clientes ocasionais em consumidores recorrentes e leais à marca.

No entanto, devido aos muitos aspectos que regem os canais de marketing e os perfis de consumo dos clientes, ter uma estratégia de oferta (offer policy) que seja capaz de disponibilizar o incentivo correto através do canal adequado e para o cliente apropriado representa um desafio considerável que afeta significativamente a performance do negócio.

Uma maneira de superar esses desafios é aplicar técnicas analíticas e de modelagem de dados para que as empresas possam mapear a probabilidade de um determinado consumidor aceitar uma oferta e assim criar uma estratégia **data-driven** ou até mesmo automatizada por um agente inteligente. Desenvolvendo perfis de clientes baseados em dados e sistemas de recomendação, é possível sugerir uma oferta personalizada para um determinado perfil de maneira a maximizar a probabilidade de consumo.

No iFood, uma das nossas abordagens a esse problema é através de algoritmos de MAB, que combinados em um sistema inteligente, são responsáveis por escolher a cada dia os descontos a serem distribuídos para determinados grupos de clientes.

# Dados disponíveis

Você terá acesso a 3 datasets:

## **offers.json:**

Esse dataset contém os ids das ofertas e metadados de cada uma delas:

- id (string) - id da oferta
- offertype (string) - o tipo da oferta. i.e. BOGO, discount, informational
- minvalue (int) - o valor mínimo que precisa ser gasto para que a oferta seja ativada
- discountvalue (int) - o valor do desconto a ser aplicado, caso a oferta seja ativada.
- channels (list of strings) - canais por onde a oferta pode ser veiculada

## **customers.json:**

Esse dataset contém atributos de cerca de 17k clientes

- age (int) - idade do cliente no momento da criação da conta
- registeredon (int) - a data em que o cliente criou a conta
- gender (string) - gênero do cliente (algumas entradas podem conter 'O' para outras opções além de M ou F)
- id (string) - id do cliente
- creditcardlimit (float) - o limite do cartão de crédito do cliente registrado no momento da criação da conta

## **transactions.json:**

Esse dataset contém cerca de 300k eventos

- event (str) - descrição do evento (transação, oferta recebida, oferta visualizada, oferta concluída, etc.).
- accountid (str) - id do cliente
- timesinceteststart (int) - tempo passado desde o começo do teste, em dias. Os dados começam em t=0
- value - (json) - pode registrar o offerid de oferta, o desconto concedido (key reward) ou o valor da transação, dependendo do tipo de evento

**Importante:** os descontos de uma oferta (key reward do campo value) são concedidos quando uma transação é realizada após a visualização de uma oferta. Nesse caso, o reward recebido será automaticamente aplicado à próxima transação. Logo, o valor da transação seguinte ao recebimento do desconto de uma oferta é o valor com o desconto já aplicado.

## O desafio

O seu objetivo é propor uma estratégia de oferta que seja automatizada por um modelo de machine learning, a fim de otimizar o ARPU (average revenue per user).

Você precisará entregar:

1. Um notebook contendo o processo de manipulação e limpeza de dados a fim de preparar um dataset unificado que será utilizado para análise e modelagem, utilizando pySpark ou python;  
> Para isso, recomendamos a criação de uma conta no databricks community (<https://community.cloud.databricks.com/>)
2. Um notebook/script contendo o processo de treino e avaliação de um modelo escolhido para ser utilizado como motor da estratégia de oferta proposta;  
> Esse passo pode ser executado utilizando as ferramentas/ linguagens com as quais você se sentir mais confortável;
3. Uma apresentação de no máximo 5 slides, destinada a uma audiência composta por líderes de negócio tecnicamente leigos, apresentando estimativas ou projeções de performance da estratégia.

## Próximos passos

Caso os materiais enviados pela pessoa candidata sejam aprovados, será realizado um convite para participar de uma entrevista técnica a respeito dos conteúdos submetidos