**Passei Direto**

**Teste - Vaga Data Scientist**

O objetivo desse teste é avaliar como você se comporta diante de um conjunto de dados e questões que queremos responder. Fique a vontade para ir além do perguntado e trazer insights extras!

Os dados para a análise estão disponíveis em anexo. Eles correspondem a uma amostra dos dados gerados pela plataforma do Passei Direto e são melhor detalhados a seguir, junto com as questões do teste.

**1ª Parte**

O Passei Direto é uma plataforma aberta (grátis) com um modelo de negócios freemium, baseado em assinaturas. É muito importante para a saúde do negócio entendermos bem o comportamento dos usuários, principalmente dos pagantes.

Nessa etapa, será usada a *BASE A*. Essa base consiste de apenas um Dataset:

* **premium\_students.json** - Registros de usuários cadastrados no Passei Direto em Novembro de 2017 e que também assinaram o Plano Premium

**Questão 1:** Dentre os usuários cadastrados em Nov/2017 que assinaram o Plano Premium, qual a probabilidade do usuário virar Premium após o cadastro em ranges de dias? A escolha dos ranges deve ser feita por você, tendo em vista os insights que podemos tirar para o negócio.

Exemplo de distribuição:

De 0 a 3 dias -> 30%

De 4 a 8 dias -> 20%

De 8 a 30 dias -> 50%

**2ª Parte**

Nessa segunda etapa, vamos fazer uma análise mais aprofundada dos nossos usuários e suas ações no Passei Direto. Será usada a *BASE B*, composta dos seguintes Datasets:

* **students.json** - Amostra de usuários que acessaram o Passei Direto no mês de Novembro de 2017. Todos os outros datasets referenciam esses usuários
* **sessions.json** - Todas as sessões dos usuários ao longo do tempo
* **subjects.json** - Disciplinas seguidas pelos usuários
* **questions.json** - Perguntas feitas pelos usuários
* **answers.json** - Respostas à perguntas feitas pelos usuários
* **fileViews.json** - Arquivos visualizados pelos usuários
* **premium\_payments.json** - Pagamentos realizados pelos usuários que assinaram o Plano Premium do Passei Direto
* **premium\_cancellations.json** - Pedidos de cancelamento do Plano Premium solicitados pelos usuários pagantes

**Questão 2:** Em Novembro de 2017 fizemos uma grande mudança no PD: o Content Restriction. Os usuários não Premium passaram a poder consumir no máximo 3 arquivos diferentes por mês. Diante dessa mudança, qual passou a ser o Lifetime Value (LTV) dos usuários Premium a partir de Novembro de 2017?

**Questão 3:** Considerando o seu conhecimento da plataforma do Passei Direto e os dados disponibilizados nesse teste, que oportunidades você vê para melhorarmos a experiência dos nossos usuários? Ofereça insights explorando o como e o porquê.

**Questão 4:** Ainda considerando os dados passados no teste, como você segmentaria os usuários da Passei Direto e quais recomendações, em termos de negócio, você forneceria para as diversas áreas do PD (considerar as áreas de Marketing, Cliente e Financeiro)? Apresentar o código de como seria feita a classificação.

**Entregáveis (enviar por email até o final do prazo combinado)**

* Código-fonte utilizado para fazer o processamento e análise dos dados
  + Idealmente colocar em um repositório no Github
* Apresentação dos resultados (formato ppt ou pdf)
  + Importante explicar também o seu processo de decisão diante das questões apresentadas
  + Vamos marcar um dia para você apresentar

Boa sorte, e nos vemos em breve.