



TESTE DE SELEÇÃO

- Engenharia de Dados



// SOBRE O OTO

O Oto é uma startup da Pmweb e braço da Wunderman Thompson Technology engajada na missão de promover a verdadeira transformação digital do varejo. Nossa plataforma transforma dados em informações relevantes para empoderar vendedores e varejistas, criando relacionamentos únicos em escala e oportunizando mais vendas.

Levamos a tecnologia ao limite, fazendo uso de volumes massivos de dados (big data), inteligência artificial, APIs e uma cultura ágil para levar ao mercado soluções visionárias e a frente do seu tempo.



// SOBRE O TESTE

Avaliar a capacidade técnica, crítica e raciocínio lógico do candidato ao construir uma aplicação e interface de usuário para resolução do problema proposto.

Receita, frequência e valor

A tabela abaixo representa as compras de alguns clientes.

Cliente	Data	Valor
1	2021-08-02	199,90
2	2022-08-01	245,20
3	2022-08-01	89,90
4	2022-08-03	545,80
2	2022-08-03	199,90
5	2022-08-04	239,90
3	2022-08-05	89,90
5	2022-08-05	119,80

Construa uma consulta (SQL) que calcule as métricas de RFV de cada um dos clientes:

- Recência = Tempo desde a última compra (dias)
- Frequência = Quantidade de compras
- Valor = Valor total em compras

Você deverá implementar o fluxo de tratamento de dados descrito abaixo, utilizando o dataset a seguir como base:

<https://oto-public.s3.amazonaws.com/natal2021.zip>

Sua tarefa:

Com o objetivo de realizar uma campanha especial de final de ano, recebemos o arquivo `natal2021.csv` do cliente. Você recebeu o desafio de realizar a limpeza deste arquivo CSV para que possamos posteriormente importá-lo em nosso banco de dados. Pensando em um cenário no qual o arquivo CSV seja muito maior e não entre na memória de uma só vez, temos os seguintes problemas para serem resolvidos:

1. Identificar o encoding do arquivo;
2. Processar o arquivo de 1000 em 1000 linhas;
3. Remover os espaços das colunas. Ex. ' Porto Alegre ' -> 'Porto Alegre';
4. Criar uma coluna CITY_ASCII no arquivo, a qual deve ser construída com base na coluna CITY. Esta coluna não pode conter acentos, minúsculas e caracteres especiais. Apenas letras, números e hífen são permitidos. Ex. 'São Paulo - abç' -> 'SAO PAULO - ABC'
5. Remover os caracteres não numéricos da coluna PHONE;
6. Salvar o arquivo .csv com a nova coluna em UTF-8;

Integração e-commerce

Um cliente gostaria de integrar os dados de seu e-commerce com nossa plataforma. Ele utiliza Vtex. ***Como você desenharia a arquitetura desta integração? Quais sistemas, fluxos e pipelines de dados você criaria?***

Considere que serão integrados os seguintes dados:

- Catálogo de produtos
- Pedidos realizados
- Cadastro de clientes

Observe que você deve apenas dissertar sobre a arquitetura, e não codificar.

O QUE SERÁ AVALIADO?

- Funcionalidade: se atinge o objetivo esperado dentro das limitações impostas.
- Legibilidade do código: será verificado se está fácil de ser lido e entendido, com nomes das funções e variáveis de acordo com a sua funcionalidade.
- Documentação, comentários e reaproveitamento de código.
- Organização na estrutura de arquivos.
- Boas práticas de desenvolvimento.
- Tempo de entrega.

ENTREGAS NECESSÁRIAS

- Você deve salvar o código fonte necessário em um repositório Git da sua escolha.
- Enviar para o e-mail: renata.rehm@otocrm.com.br, everton.cunha@otocrm.com.br, tobias.jacoby@otocrm.com.br