

Desafio Engenheiro de Dados

O desafio consiste em criar um *snippet* de código e responder uma pergunta sobre como escalar essa solução. A disponibilização do *snippet* e resposta da pergunta podem ser feitas via algum site que permita o compartilhamento de código, GitHub ou até mesmo pelo e-mail. Abaixo seguem algumas sugestões de sites para criar os *snippets*.

- Scala - <https://scastie.scala-lang.org/>
- Python e Java - <https://www.onlinegdb.com/>

Sinta-se livre para usar eles ou outro de sua preferência.

O Problema

O desafio consiste em calcular o ganho total da empresa *Acquirer LTDA*, que o obtêm na prestação do serviço de locação de máquinas de cartão de crédito para seus clientes. Esse ganho é calculado sobre um percentual das transações de cartão de crédito realizadas por eles. O cálculo é feito baseado em dois conjuntos de dados, transações e contratos.

Transações

transaction_id	client_id	total_amount	discount_percentage
1	3545	3000	6.99
2	3545	4500	0.45
3	3509	69998	0
4	3510	1	null
5	4510	34	40

Tabela com todas as transações das máquinas de cartão de crédito realizadas pelos clientes da *Acquirer LTDA*.

Contratos

contract_id	client_id	client_name	percentage	is_active
3	3545	Magazine Luana	2.00	true
4	3545	Magazine Luana	1.95	false
5	3509	Lojas Italianas	1	true
6	3510	Carrerfive	3.00	true

Tabela com os detalhes financeiros dos contratos da *Acquirer LTDA* com seus clientes.

O resultado esperado é o ganho total da *Acquirer LTDA*, que para o exemplo é: 845,411.

O cálculo deve ser feito cruzando os dados de transações com contratos (considerando apenas o contrato ativo do cliente). Para se obter o ganho total, é necessário multiplicar o percentual do contrato pelo valor líquido da transação, ou seja, já descontando o valor do desconto aplicado na transação.

A Pergunta

Além do código acima, considere que uma escala de ~200 milhões de transações por dia e que o cálculo deverá apresentar um resultado do valor total do mês. Descreva em até 500 palavras que tecnologias e arquitetura você usaria para escalar a solução acima.

Em caso de dúvidas, pergunte.