



Fabio Kfourri <kfourri.fabio@gmail.com>

Desafio Área de dados Dotz | Tecer Recursos + Humanos | Deadline Terça 20.10

1 mensagem

Tamara Costa <tamara.costa@tecer.us>

15 de outubro de 2020 22:37

Para: kfourri.fabio@gmail.com

Cc: Josimar dos Santos Junior <josimar.junior@dotz.com>

Olá **Fabio!**

Obrigada pelo interesse em fazer parte do time **Dotz**! Esse é um teste com o objetivo de conhecer um pouco mais sobre a sua forma de trabalhar com problemas que envolvem Engenharia de dados.

O problema:

O presente problema se refere aos dados de uma empresa que produz e vende máquinas industriais. O objetivo deste desafio é

1. Fazer a modelagem conceitual dos dados
2. Criar a infraestrutura necessária
3. Criar todos os artefatos necessários para carregar os arquivos para o banco criado.

Os seguintes arquivos devem ser importados para o banco de dados:

- price_quote.csv

Esse arquivo contém informações sobre price quotes dos fornecedores. Os preços podem ser cotados de 2 formas:

1. Bracket pricing: possui múltiplos níveis de compra, baseado em quantidades (isso significa que o custo é dado assumindo-se a compra de certa quantidade de tubos)
2. Non-bracket pricing: possui uma quantidade mínima a ser pedida (min_order) para que o preço seja aplicado.

Cada uma dessas cotações é emitida com uma annual_usage, que se refere a estimativa de quantos conjuntos de tubos serão comprados em determinado ano.

- Bill_of_materials.csv

Esse arquivo contém a lista de componentes e as quantidades que foram utilizadas em cada conjunto de tubos.

- comp_boss.csv

Esse arquivo contém as informações sobre cada componente.

Resolução:

Para a realização deste teste você sugerimos que utilize a nuvem do **Google (GCP)**, **Data Flow (GCP)** e **Python**. Pedimos também que compartilhe conosco o código fonte (caso queira utilizar um repositório git, sinta-se à vontade). Aguardamos um processo de ingestão de dados e uma visualização (relatório) consumindo os dados ingeridos pelo seu processo.

Além disso, inclua um arquivo README.md onde você deve compartilhar as decisões de arquitetura, implementação e instruções sobre como executar o software.

Descrição dos dados

<https://drive.google.com/drive/folders/1CEjgMt0F5qsftNIUCefYafyIWLLSMFE3?usp=sharing>**Prazo para finalização do desafio: 20/10 às 10h00**

Responder esse e-mail copiando o Josimar (também em cópia) e informar Tamara via whats app.

Bom desafio! :)

Abraços,

tecer

Tamara Costa

Cofundadora | Caçadora de talentos

📞 11 95780-9381

in



Esta mensagem de e-mail, juntamente com seus anexos, se houver, é confidencial e pode conter informações sujeitas a privilégios legais. Se você não é o destinatário pretendido ou recebeu este e-mail por engano, informe-nos imediatamente e exclua esta mensagem. Qualquer cópia não autorizada desta mensagem (e anexos) ou distribuição não autorizada das informações aqui contidas é proibida.