

Teste Engenheiro de Dados Esfera

1 - Construção do ETL

Neste teste, você será responsável por desenvolver um processo de ETL para processar arquivos CSV e armazenar os dados em um banco de dados, conforme especificado. Os dados disponíveis representam o orçamento do Estado de São Paulo de 2019. Os requisitos são os seguintes:

- Adequar os tipos de dados para os mais apropriados.
- Utilizar Python e/ou SQL para processamento dos dados.
- Hospedar o código do projeto e fornecer uma explicação detalhada no GitHub.

Entregáveis:

Os dados tratados e em tabelas no banco de dados.

2 - Orquestração de Tarefas (Não obrigatório, mas é um diferencial)

O objetivo deste exercício é avaliar o conhecimento em automação do fluxo de dados.

O candidato tem liberdade para demonstrar seus conhecimentos em orquestração, infraestrutura e containerização do projeto utilizando Docker.

Entregáveis:

A ferramenta de orquestração deve incluir cada etapa do ETL, sendo preferencialmente Airflow.

3 – Visualização e análise de dados

Criar um painel em uma ferramenta de dataviz para responder quais são as 5 maiores fontes de recursos e os 5 maiores tipos de despesa.

O candidato tem liberdade para demonstrar insights que achar pertinente.