Teste Prático – Analista de Business Intelligence Pleno.

Janeiro – 2025

Versão: Candidato

**Contexto**

A Neotass Marketing firmou parceria com um cliente no segmento B2B para desenvolver uma solução que potencialize a análise de dados de oportunidades de negócio e vendas. Este cliente, que atua como parceiro de diversas empresas, registra no sistema informações sobre oportunidades de negócio que seus próprios parceiros submetem. Além disso, o cliente disponibiliza dados de vendas realizadas (sellout) para análise de desempenho.

O objetivo deste projeto é consolidar, transformar e estruturar os dados recebidos de diferentes fontes em um formato integrado e limpo, possibilitando análises avançadas e a geração de insights estratégicos.

**Como responsável por este projeto, você deverá:**

1. **Criar um ETL (Extract, Transform, Load) usando a linguagem Python para processar os dados.**
2. **Integrar os dados processados em um banco de dados dimensional, projetado para facilitar consultas analíticas. (Podendo ser Fictício em formato .xlsx ou .parquet)**
3. **Desenvolver Documentação das etapas do processo de dados do projeto.**

**Bases de Dados Disponíveis**

1. **registros\_oportunidades.json**Esta base contém informações sobre oportunidades de negócio registradas no sistema do cliente. **Detalhes da base:**
   * Inclui informações como IDs de oportunidades, parceiros responsáveis pelo registro, descrição, datas e status das oportunidades etc.
2. **sellout.parquet**Esta base inclui dados detalhados das vendas realizadas (sellout).  
   **Detalhes da base:**
   * Contém informações como IDs de vendas, IDs dos parceiros, produtos vendidos, quantidades, valores e datas de venda.

**Requisito do Projeto:**

1. O código deve ser desenvolvido obrigatoriamente em Python, utilizando boas práticas de programação. Sinta-se à vontade nas escolhas das bibliotecas a serem utilizadas.
2. O ETL deve ser capaz de organizar os dados no modelo de banco de dados dimensional podendo ser o banco fictício em formado .xlsx. Desde que seguido a seguinte estrutura para as tabelas fatos:
   1. fato\_registro\_oportunidade: id\_oportunidade, id\_parceiro, id\_produto, data\_registro, quantidade, valor\_total, status.
   2. fato\_sellout: id\_sellout, id\_parceiro, id\_produto, data\_fatura, nf, quantidade, valor\_total.

O projeto deve ser enviado para a **Neotass Marketing** no prazo de **[inserir número de dias] dias** a partir do recebimento deste teste.

Embora não seja obrigatório, será considerado um **diferencial** caso o candidato:

1. Desenvolva o **ETL utilizando o orquestrador Airflow**, demonstrando sua capacidade de trabalhar com pipelines de dados automatizados.
2. Estruture o banco de dados no **MySQL** e envie o script SQL para a criação das tabelas (.sql) juntamente com o restante do projeto.

**Entregáveis:**

1. Link do projeto publicado no Github.
2. **Dados Processados**: Arquivos no formato .xlsx ou .parquet após a etapa de transformação.
3. **Banco de Dados**:
   1. Caso opte por MySQL, inclua o arquivo .sql com a criação das tabelas e exemplos de inserções.
   2. Caso utilize outro banco, forneça instruções claras para acesso.
4. **Relatório Explicativo**: Documento em .pdf ou .docx detalhando as etapas do ETL, desafios enfrentados e decisões tomadas.
5. **Orquestração (Opcional)**: Caso utilize o Airflow, inclua os arquivos de DAGs e instruções para execução.

Desejamos boa sorte no desenvolvimento do Projeto.

Neotass is a Journey!