

Avaliação de Conhecimento em Engenharia

Avaliação:

Cada um dos pontos listados em “o que será avaliado”, receberá uma das seguintes classificações: “não atendeu”, “atendeu parcialmente”, “atendeu plenamente”, “superou as expectativas”.

Material fornecido:

- Desafios 1 a 3: três arquivos CSV das tabelas “entidade”, “pessoa” e “transação”;
- Descritivo das tabelas.

Desafio 1 - Importação dos dados

Objetivo: definir e construir uma pipeline para importação da massa de dados disponibilizada.

O que fazer: Utilize os recursos que achar necessários para importar os arquivos disponibilizados em uma base de dados (PostgreSQL, MySQL)

O que deve ser entregue: código fonte do processo ETL criado, sugere-se a utilização python/spark/pyspark.

O que será avaliado:

- Domínio do processo ETL para importação dos dados
- Sanitização e processamento das informações;
- Tratamento para a coluna cod_cooperativa: 4 dígitos. Exemplo: cod_cooperativa = "123", deve-se ajustar para "0123"
- Aplicação correta e uso de PK's

Desafio 2 - Criação de tabela “Flat”

Objetivo: gerar uma tabela “flat” com o objetivo de disponibilizar informações em um relatório.

O que fazer: criar uma tabela única contendo informações dos associados, agregando as informações abaixo:

- Associado Frequente: associado transacionou uma ou mais vezes em cada um dos últimos 3 meses, em qualquer modalidade.
- Associado Ativo no Crédito: associado transacionou nos últimos 3 meses na modalidade Crédito.
- Associado Ativo no Débito: associado transacionou nos últimos 3 meses na modalidade Débito.
- Utilize SQL para a construção desta tabela e geração destes indicadores

O que deve ser entregue: códigos e scripts utilizados na construção da tabela e manipulação dos dados.

O que será avaliado:

- Domínio do processo ETL para o processamento dos dados.
- Domínio da linguagem Sql para consulta e manipulação dos dados.

Desafio 3 - Modelagem Fato X Dimensão

Objetivo: definir e construir uma modelagem Fato X Dimensão a partir dos dados disponibilizados.

O que fazer: utilizando as tabelas de exemplos informadas, crie uma modelagem Fato/Dimensão, criando as estruturas que julgar necessárias.

O que deve ser entregue:

- Diagrama ER do modelo dimensional aplicado
- Códigos fontes utilizados para a construção e persistência dos dados das tabelas

O que será avaliado:

- Correta representação do modelo Dimensional;
- Uso adequado de granularidade;
- Aplicação correta da Cardinalidade (PK, FK).