

**Caio Martim Barros**  
**Case Técnico Estágio Engenharia de Dados**  
**EloGroup**

1. Apresente as características do modelo de dados fornecido, destacando sua visão geral acerca do modelo de dados e tecendo críticas e comentários.;

**Resposta:**

Esse modelo de dados proposto é classificado como um modelo multidimensional, a qual é aclamado como um star-schema. Dele, conseguimos analisar que é composto por dois tipos de tabelas, uma principal chamada de 'fact table' dado pelo nome no modelo de dados: 'FT\_vendas' e o outro tipo de tabela chamada de 'dimension table', que são as tabelas restantes desse modelo.

Esse star-schema possui uma tabela fato FT\_Vendas que contém as chaves estrangeiras que estarão relacionadas com as tabelas dimensões, que possuem as chaves primárias para fazer essa relação. Portanto, a tabela FT\_Vendas armazena as medidas que serão analisadas de acordo com as dimensões que queremos realizar, interagindo, assim, a tabela fato com o restante das tabelas com simples metodologias como um inner join para esse acesso.

Nesse contexto, no ponto de vista geral, o star-schema é um modelo importante para a análise de dados e fácil de ser manipulado e lido, esse esquema possui características importantes para a realização de análises com propósito de visualização por períodos, se tratando com dados não normalizados e está relacionado com os sistemas OLAP. Além disso, esse modelo de dados garante uma boa performance pelo fato das queries serem simples para serem carregadas e são comumente usadas em serviços de análise de dados, como PowerBI e pode ser usado também na construção de Data Warehouses e Data Mart.

Todavia, o star-schema atua no formato 'one-to-many', a qual a tabela fato se relaciona com o restante das tabelas dimensões, então para uma arquitetura do formato 'many-to-many', ela não é usada. Nesse âmbito, outra crítica seria no conceito desse esquema usar dados não normalizados, o que indica em uma limitação da análise dos dados, não permitindo visualizações mais complexas, além de que se tratando de sistema OLAP, não se pode adicionar ou atualizar os dados pois pode comprometer a análise, dessa forma, o star-schema é usado mais na metodologia 'read-only'.

2. Demonstre seu conhecimento em SQL, inserindo código abaixo de cada solicitação:

(2.1) Quantidade vendida no período entre 01/04/2020 a 01/04/2021;

```
1 SELECT
2 Dia,
3 Mes,
4 Ano,
5 Qtde_vendida
6 FROM FT_Vendas
7 INNER JOIN DIM_Tempo
8     ON DIM_Tempo.ID_tempo = FT_Vendas.vendas_tempo
9     WHERE Dia = 1 AND Mes = 4 AND (Ano = 2020 OR Ano = 2021);
```

(2.2) Quantidade de clientes únicos do sexo feminino e estado civil divorciada;

```
1 SELECT
2 COUNT(*) AS quantidade_clientes
3 FROM DIM_Cliente
4 WHERE Estado_civil = "Divorciada" AND Sexo = "Feminino";
```

(2.3) Buscar os produtos com código 55, 120, 142 e que tiveram receita maior que R\$ 120,00;

```
1  SELECT
2  ID_produto ,
3  Cod_produto ,
4  Receita_venda
5  FROM FT_Vendas
6  INNER JOIN DIM_Produto
7      ON DIM_Produto.ID_produto = FT_Vendas.vendas_produto
8      WHERE (DIM_Produto.Cod_produto = 55 OR DIM_Produto.Cod_produto = 120
9      OR DIM_Produto.Cod_produto = 142) AND Receita_venda > 120;
```

(2.4) Verificar a receita da loja com o código 32 no período igual ou superior a 01/04/2021;

```
1  SELECT
2  Dia ,
3  Mes ,
4  Ano ,
5  Receita_venda
6  FROM FT_Vendas
7  INNER JOIN DIM_Loja
8      ON DIM_Loja.ID_loja = FT_Vendas.vendas_loja
9      WHERE DIM_Loja.ID_loja = 32 AND Dia >= 1 AND Mes >= 4 AND Ano >= 2021
10 GROUP BY DIM_Loja.ID_loja;
```

**3.** Por fim, descreva o que você melhoraria nesse modelo;