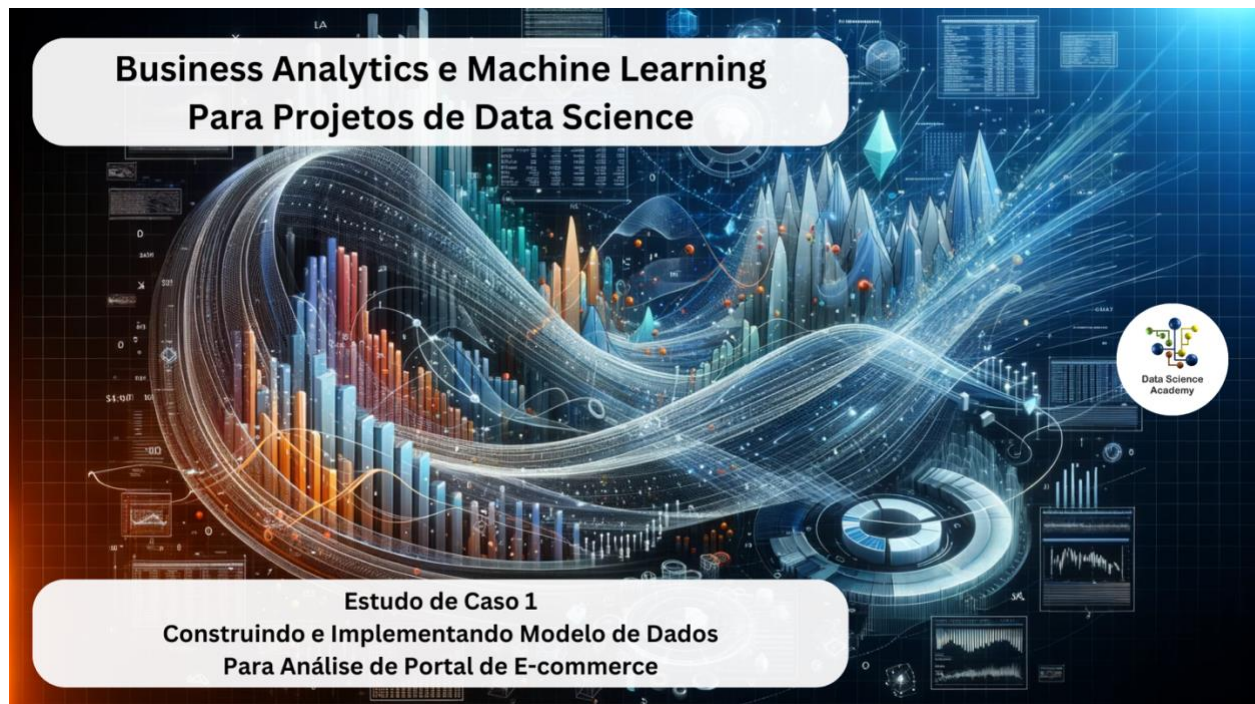




Business Analytics e Machine Learning Para Projetos de Data Science

Estudo de Caso 1 Modelo Dimensional



Business Analytics e Machine Learning Para Projetos de Data Science

Estudo de Caso 1 Construindo e Implementando Modelo de Dados Para Análise de Portal de E-commerce

O modelo dimensional é centrado em torno de tabelas de fatos, que são ligadas a várias tabelas de dimensões.

O modelo dimensional é dividido em dois tipos de tabelas:

Tabela Fato: A tabela fato armazena as métricas quantitativas (como quantidade vendida e faturamento) e as chaves estrangeiras para as tabelas de dimensões. Você pode ter mais de uma tabela fato no mesmo modelo.

Tabelas de Dimensão: Cada dimensão (Produto, Cliente, Data, Distribuidor) contém atributos descritivos. Por exemplo, a dimensão Data pode ser usada para análises de tendência ao longo do tempo.

O modelo dimensional tem as seguintes características:

Desnormalização: Em modelos dimensionais, as tabelas são frequentemente desnormalizadas para melhorar o desempenho das consultas. Ou seja, há maior agrupamento dos dados.

Análise: Este modelo facilita análises como a performance de vendas por categoria de produto, análise do comportamento de compra dos clientes e tendências de vendas ao longo do tempo.

Lembre-se do objetivo deste projeto: **Você foi contratado pela DSA10 para construir um Data Warehouse que permita extrair métricas de vendas, de operação e de comportamento de compra dos clientes.**

Aqui está a proposta de Modelo Dimensional para o DW da DSA10:

Tabela Fato: Vendas

- ClientelID (chave estrangeira para Dimensão Cliente)
- ProdutoID (chave estrangeira para Dimensão Produto)
- DistribuidorID (chave estrangeira para Dimensão Distribuidor)
- DataID (chave estrangeira para Dimensão Data)
- Quantidade Vendida
- Faturamento
- Custo Frete

Dimensão Cliente

- ClientelID (chave primária)
- Nome
- Endereço
- Cidade do Cliente
- País do Cliente

Dimensão Produto

- ProdutoID (chave primária)
- Nome
- Descrição
- Preço
- Nome Categoria

Dimensão Distribuidor

- DistribuidorID (chave primária)
- Nome do Distribuidor
- Cidade do Distribuidor
- País do Distribuidor

Dimensão Data

- DataID (chave primária)
- Data
- Dia
- Mês
- Ano
- Dia da Semana

Dimensão Avaliação (opcional, dependendo das necessidades de análise)

- AvaliaçãoID (chave primária)
- ProdutoID
- ClienteID
- Pontuação
- Comentário

Dimensão Categoria (opcional, dependendo das necessidades de análise)

- CategoriaID (chave primária)
- Nome Categoria
- Tipo de Categoria

Este modelo dimensional oferece uma estrutura eficaz para armazenar e analisar dados em um ambiente de Data Warehouse, permitindo consultas rápidas e eficientes, essenciais para a análise de negócios e tomada de decisões no portal de e-commerce.

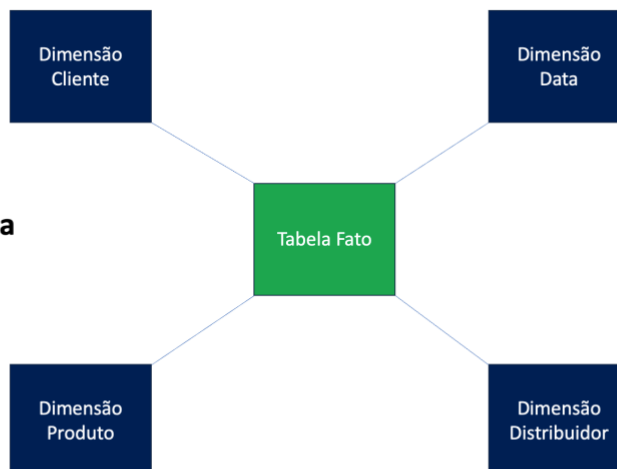
Este é o diagrama do modelo dimensional no formato Star Schema (o qual iremos implementar na prática):



Business Analytics e Machine Learning Para Projetos de Data Science



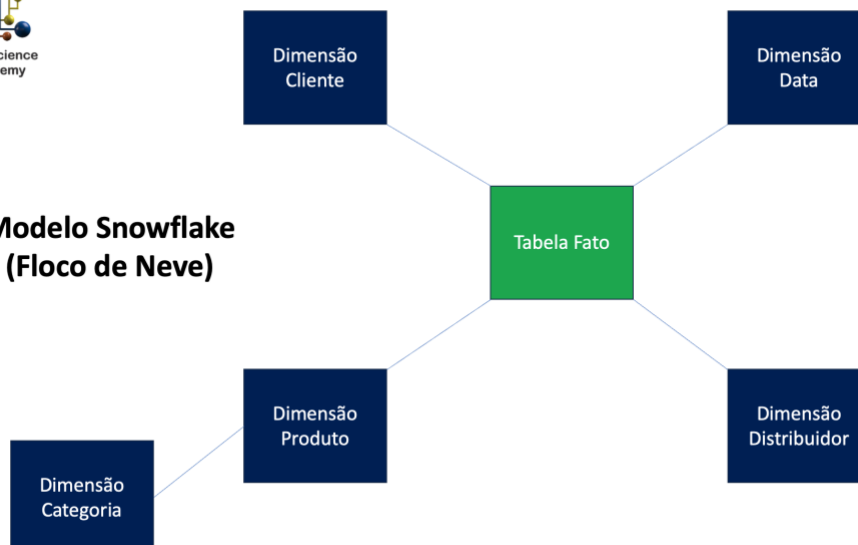
Modelo Star Schema (Esquema Estrela)



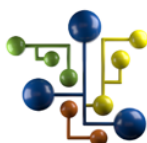
Este é o diagrama modelo dimensional no formato Snowflake (opcional):



Modelo Snowflake (Floco de Neve)



Podemos agora construir os modelos lógico e físico a partir desse modelo dimensional.



Equipe DSA

Muito Obrigado!
Continue Trilhando Uma Excelente Jornada de Aprendizagem.