**Centro Universitário Maurício de Nassau**

THAIS HELENA RAMOS DE MELO

**MODELAGEM DIMENSIONAL**

Recife/PE

2024

Thais Helena Ramos de Melo

Matrícula: 01068175

**MODELAGEM DIMENSIONAL**

Artigo apresentado ao Curso de Ciência da Computação, Turma: 3NA, do Centro Universitário Uninassau, como requisito parcial para a disciplina de Database Application.

Orientador: Prof. Filipe Nascimento.

Recife/PE

2024

**Sumário**

[**1 Introdução** 5](#_Toc167746350)

[**2 Star Schema** 6](#_Toc167746351)

[**2.1 Principais características** 6](#_Toc167746352)

[**2.1.1 Simplicidade** 6](#_Toc167746353)

[**2.1.2 Desempenho eficiente** 6](#_Toc167746354)

[**2.1.3 Facilidade de consulta** 6](#_Toc167746355)

[**2.1.4 Flexibilidade** 6](#_Toc167746356)

[**2.1.5 Suporte a consultas complexas** 7](#_Toc167746357)

[**2.1.6 Facilidade de manutenção** 7](#_Toc167746358)

[**2.2 Casos de uso** 7](#_Toc167746359)

[**2.2.1 Análise de vendas e receitas** 7](#_Toc167746360)

[**2.2.2 Monitoramento de desempenho operacional** 7](#_Toc167746361)

[**2.2.3 Análise de marketing e campanhas** 7](#_Toc167746362)

[**2.2.4 Relatórios financeiros e análise de custos** 8](#_Toc167746363)

[**3 Snowflake schema** 9](#_Toc167746364)

[**3.1 Principais Características** 9](#_Toc167746365)

[**3.1.1 Normalização de dados** 9](#_Toc167746366)

[**3.1.2 Economia de espaço** 9](#_Toc167746367)

[**3.1.3 Integridade dos dados** 9](#_Toc167746368)

[**3.1.4 Flexibilidade** 9](#_Toc167746369)

[**3.1.5 Desempenho de consulta** 10](#_Toc167746370)

[**3.1.6 Manutenção e gerenciamento** 10](#_Toc167746371)

[**3.2 Casos de uso** 10](#_Toc167746372)

[**3.2.1 Ambientes com grande volume de dados** 10](#_Toc167746373)

[**3.2.2 Aplicações que exigem integridade de dados** 10](#_Toc167746374)

[**3.2.3 Análise multidimensional complexa** 10](#_Toc167746375)

[**3.2.4 Sistemas de data warehousing corporativos** 11](#_Toc167746376)

[**3.2.5 Aplicações que requerem flexibilidade e evolução** 11](#_Toc167746377)

[**4 Principais diferenças entre os Sistemas Star Schema e Snowflake** 12](#_Toc167746378)

[**4.1 Estrutura** 12](#_Toc167746379)

[**4.2 Redundância de Dados** 12](#_Toc167746380)

[**4.3 Desempenho** 12](#_Toc167746381)

[**4.4 Complexidade e Flexibilidade** 12](#_Toc167746382)

[**4.5 Aplicações e Casos de Uso** 12](#_Toc167746383)

[**5 Implementação de um exemplo de Modelo Star Schema** 14](#_Toc167746384)

[**5.1 Modelo Lógico** 14](#_Toc167746385)

[**5.2 Modelo Físico** 14](#_Toc167746386)

[**6 Cubos OLAP** 17](#_Toc167746387)

[**6.1 Arquitetura** 17](#_Toc167746388)

[**6.2 Implementação** 18](#_Toc167746389)

[**6.3 Casos de uso** 19](#_Toc167746390)

[**6.3.1 Análise de Vendas e Desempenho Financeiro** 19](#_Toc167746391)

[**6.3.2 Gestão de Estoque e Cadeia de Suprimentos** 19](#_Toc167746392)

[**6.3.3 Análise de Clientes e Segmentação de Mercado** 19](#_Toc167746393)

[**6.3.4 Planejamento e Orçamento Empresarial** 20](#_Toc167746394)

[**6.3.5 Análise de Recursos Humanos** 20](#_Toc167746395)

[**6.3.6 Análise de Dados de Saúde** 20](#_Toc167746396)

[**7 Implementação de um Exemplo de Cubo OLAP no MySQL** 21](#_Toc167746397)

[**7.1 Consulta** 21](#_Toc167746398)

[**7.2 Resultado** 22](#_Toc167746399)

[**7.3 Consulta** 22](#_Toc167746400)

[**7.4 Resultado** 22](#_Toc167746401)

[**Referências Bibliográficas** 23](#_Toc167746402)

# **1 Introdução**

Nos ambientes de Business Intelligence (BI) e análise de dados, a eficácia da tomada de decisões depende diretamente da qualidade e da estrutura dos dados disponíveis. Duas metodologias amplamente empregadas na organização e na análise de dados são os esquemas dimensionais Star Schema e Snowflake Schema. Ambos desempenham um papel importante na organização e na acessibilidade dos dados para análises complexas e abrangentes.

Neste estudo, vamos explorar o conceito, as características, os casos de uso e a implementação prática dos esquemas Star Schema e Snowflake Schema, assim como dos Cubos OLAP. Além disso, vamos destacar as principais diferenças entre os esquemas Star e Snowflake e discutir a implementação de um exemplo de modelo de cada um desses esquemas.

# **2 Star Schema**

O Star Schema, Esquema em Estrela, é um modelo de design de banco de dados utilizado em data warehousing e business intelligence (BI). Ele consiste em uma tabela central, conhecida como tabela de fatos, que contém métricas ou medidas de negócio, cercada por várias tabelas de dimensão que descrevem os contextos em torno dessas medidas. Essas tabelas de dimensão geralmente são desnormalizadas, o que significa que elas contêm redundância de dados para facilitar consultas mais rápidas e simples.

Na estrutura do Star Schema, a tabela de fatos é conectada às tabelas de dimensão por chaves estrangeiras. Cada chave estrangeira na tabela de fatos corresponde a uma chave primária em uma tabela de dimensão, estabelecendo assim as relações entre os dados. As tabelas de dimensão fornecem informações adicionais sobre os dados na tabela de fatos, permitindo uma análise mais detalhada e contextualizada.

## **2.1 Principais características**

### **2.1.1 Simplicidade**

Conhecido por sua estrutura simples e fácil compreensão. Ele consiste em uma tabela central de fatos e várias tabelas de dimensão, formando uma estrutura de estrela que é intuitiva para analistas e usuários finais.

### **2.1.2 Desempenho eficiente**

Devido à sua estrutura desnormalizada, o Star Schema oferece um desempenho eficiente para consultas de agregação e análises dimensionais. A redundância de dados nas tabelas de dimensão facilita consultas rápidas e simplificadas, resultando em tempos de resposta mais curtos.

### **2.1.3 Facilidade de consulta**

Simplifica o processo de consulta, permitindo que os usuários executem facilmente análises multidimensionais. As relações entre a tabela de fatos e as tabelas de dimensão são claras e diretas, facilitando a navegação e a exploração dos dados.

### **2.1.4 Flexibilidade**

É flexível e pode ser facilmente adaptado às necessidades específicas de negócios. Novas dimensões ou medidas podem ser adicionadas sem afetar a estrutura principal, proporcionando escalabilidade e adaptabilidade ao ambiente de data warehousing.

### **2.1.5 Suporte a consultas complexas**

Apesar de sua simplicidade, é capaz de lidar com consultas complexas que envolvem múltiplas dimensões e agregações. Ele fornece uma base sólida para análises avançadas e relatórios detalhados.

### **2.1.6 Facilidade de manutenção**

Devido à sua estrutura simplificada e desnormalizada, o Star Schema é mais fácil de manter e gerenciar em comparação com outros modelos mais complexos. Isso facilita a administração do sistema e reduz os custos operacionais.

Essas características fazem do Star Schema uma escolha popular para projetos de data warehousing e business intelligence, especialmente em ambientes onde a simplicidade, o desempenho e a facilidade de uso são prioridades.

## **2.2 Casos de uso**

### **2.2.1 Análise de vendas e receitas**

Frequentemente utilizado em empresas para analisar vendas e receitas. A tabela de fatos pode conter métricas como quantidade vendida, receita bruta e lucro, enquanto as tabelas de dimensão podem incluir informações sobre produtos, clientes, localizações e datas. Isso permite análises detalhadas sobre o desempenho de vendas, segmentação de clientes e tendências de mercado.

### **2.2.2 Monitoramento de desempenho operacional**

Eficaz para monitorar o desempenho operacional de uma empresa. Métricas de desempenho, como tempo de produção, eficiência de equipamentos e taxa de erro, podem ser armazenadas na tabela de fatos, enquanto as tabelas de dimensão fornecem contexto, como informações sobre linhas de produção, funcionários e turnos. Isso permite uma análise aprofundada do desempenho operacional e identificação de áreas de melhoria.

### **2.2.3 Análise de marketing e campanhas**

Útil para analisar o desempenho de campanhas de marketing e publicidade. Métricas como taxa de cliques, conversões e retorno sobre investimento (ROI) podem ser armazenadas na tabela de fatos, enquanto as tabelas de dimensão podem conter informações sobre canais de marketing, segmentos de público-alvo e datas de campanha. Isso possibilita avaliar a eficácia de diferentes estratégias de marketing e otimizar os gastos com publicidade.

### **2.2.4 Relatórios financeiros e análise de custos**

Amplamente utilizado para análise financeira e de custos em empresas. Métricas financeiras, como receita, despesas e margens de lucro, podem ser armazenadas na tabela de fatos, enquanto as tabelas de dimensão podem incluir informações sobre contas contábeis, centros de custo e períodos fiscais. Isso permite a criação de relatórios financeiros detalhados e análise de tendências de desempenho financeiro ao longo do tempo.

# **3 Snowflake schema**

O Snowflake Schema, Esquema de Floco de Neve, é um modelo de design de banco de dados utilizado em data warehousing e business intelligence (BI). Ele é uma extensão do modelo Star Schema e visa normalizar ainda mais a estrutura das tabelas de dimensão, eliminando redundâncias de dados e economizando espaço de armazenamento.

Na estrutura do Snowflake Schema, as tabelas de dimensão são organizadas em uma estrutura hierárquica de várias tabelas normalizadas, assemelhando-se a um floco de neve, daí o nome. Isso significa que as dimensões são divididas em subdimensões, e cada subdimensão é armazenada em uma tabela separada, conectada por chaves estrangeiras. Por exemplo, uma dimensão de tempo pode ser normalizada em várias tabelas, como ano, mês e dia.

## **3.1 Principais Características**

### **3.1.1 Normalização de dados**

Normaliza ainda mais a estrutura das tabelas de dimensão em comparação com o modelo Star Schema. Isso significa que as dimensões são divididas em subdimensões e armazenadas em várias tabelas separadas, reduzindo a redundância de dados e economizando espaço de armazenamento.

### **3.1.2 Economia de espaço**

Devido à sua estrutura normalizada, é mais eficiente em termos de espaço de armazenamento em comparação com o Star Schema. A eliminação de redundâncias de dados resulta em uma utilização mais eficiente do armazenamento, o que é especialmente importante em ambientes que lidam com grandes volumes de dados.

### **3.1.3 Integridade dos dados**

A normalização adicional ajuda a manter a integridade dos dados, reduzindo a probabilidade de inconsistências e duplicações. Isso é especialmente importante em ambientes onde a precisão dos dados é crucial para análises precisas e tomadas de decisão informadas.

### **3.1.4 Flexibilidade**

Apesar de sua estrutura mais normalizada, ainda oferece flexibilidade para adicionar novas dimensões e subdimensões conforme necessário. Isso permite que as organizações adaptem o esquema às suas necessidades específicas de análise e evoluam ao longo do tempo.

### **3.1.5 Desempenho de consulta**

Embora a normalização adicional possa tornar as consultas mais complexas devido à necessidade de junções entre tabelas, ainda pode oferecer um bom desempenho de consulta quando adequadamente otimizado. Índices e estratégias de modelagem adequadas podem ajudar a mitigar quaisquer impactos negativos no desempenho.

### **3.1.6 Manutenção e gerenciamento**

A estrutura mais normalizada pode exigir um pouco mais de esforço na manutenção e gerenciamento em comparação com o Star Schema. No entanto, com boas práticas de modelagem e administração de banco de dados, é possível garantir uma operação eficiente e confiável do esquema.

## **3.2 Casos de uso**

### **3.2.1 Ambientes com grande volume de dados**

Especialmente útil em ambientes que lidam com grandes volumes de dados, onde a economia de espaço de armazenamento é crucial. A normalização adicional ajuda a reduzir a redundância de dados, permitindo uma utilização mais eficiente do armazenamento e facilitando a escalabilidade do sistema.

### **3.2.2 Aplicações que exigem integridade de dados**

Em cenários onde a integridade dos dados é de grande importância, como em sistemas financeiros ou de saúde, pode ser uma escolha adequada. A normalização ajuda a garantir a consistência e a precisão dos dados, reduzindo a probabilidade de inconsistências e duplicações.

### **3.2.3 Análise multidimensional complexa**

Embora o Snowflake Schema possa tornar as consultas mais complexas devido à necessidade de junções entre tabelas, ele ainda é adequado para análises multidimensionais complexas. Em ambientes onde análises detalhadas e contextuais são necessárias, pode fornecer uma estrutura flexível e organizada para explorar os dados em diferentes perspectivas.

### **3.2.4 Sistemas de data warehousing corporativos**

É comumente utilizado em projetos de data warehousing corporativos, onde a economia de espaço de armazenamento e a integridade dos dados são prioridades. Ele oferece uma estrutura mais normalizada e organizada para armazenar e analisar dados de várias fontes, fornecendo uma base sólida para tomada de decisões baseada em dados.

### **3.2.5 Aplicações que requerem flexibilidade e evolução**

Embora o Snowflake Schema seja mais normalizado do que o Star Schema, ainda oferece flexibilidade para adicionar novas dimensões ou subdimensões conforme necessário. Isso o torna adequado para aplicações que precisam evoluir ao longo do tempo e se adaptar às mudanças nos requisitos de negócios e nas fontes de dados.

# **4 Principais diferenças entre os Sistemas Star Schema e Snowflake**

## **4.1 Estrutura**

- Star Schema: As tabelas de dimensão são desnormalizadas, o que significa que elas contêm todos os atributos relevantes em uma única tabela, reduzindo a necessidade de junções entre tabelas.

- Snowflake Schema: As tabelas de dimensão são normalizadas ainda mais, o que resulta em uma estrutura hierárquica de várias tabelas, onde cada subdimensão é armazenada em uma tabela separada.

## **4.2 Redundância de Dados**

- Star Schema: Permite uma certa quantidade de redundância de dados nas tabelas de dimensão, o que pode resultar em uma utilização mais eficiente do armazenamento e um desempenho de consulta mais rápido.

- Snowflake Schema: Busca reduzir a redundância de dados ao normalizar ainda mais as tabelas de dimensão, o que economiza espaço de armazenamento, mas pode aumentar a complexidade das consultas devido à necessidade de junções entre tabelas.

## **4.3 Desempenho**

- Star Schema: Devido à sua estrutura desnormalizada, geralmente oferece um desempenho de consulta mais rápido para consultas simples e agregações.

- Snowflake Schema: Pode apresentar um desempenho de consulta um pouco mais lento devido à necessidade de junções entre tabelas, especialmente em consultas que envolvem várias subdimensões.

## **4.4 Complexidade e Flexibilidade**

- Star Schema: É mais simples e fácil de entender, oferecendo uma abordagem direta para análises multidimensionais. No entanto, pode ser menos flexível em termos de evolução e adição de novas dimensões.

- Snowflake Schema: É mais complexo devido à sua estrutura normalizada, mas oferece mais flexibilidade para adicionar novas dimensões ou subdimensões conforme necessário.

## **4.5 Aplicações e Casos de Uso**

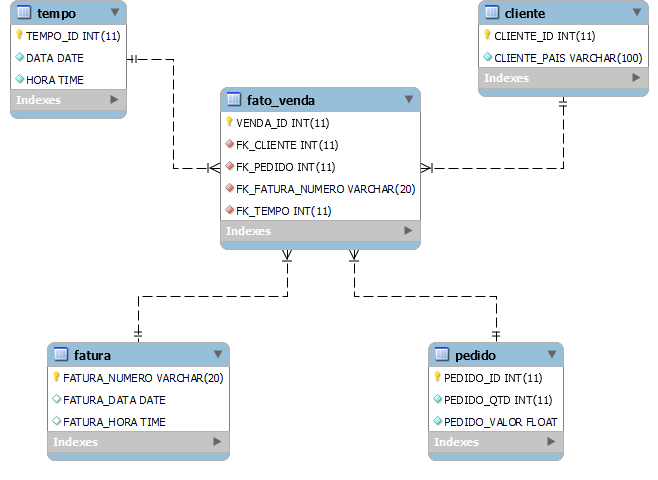
- Star Schema: É frequentemente usado em ambientes onde a simplicidade, o desempenho de consulta e a facilidade de uso são prioridades, como em análises de vendas, marketing e desempenho operacional.

- Snowflake Schema: É mais adequado para ambientes onde a economia de espaço de armazenamento e a integridade dos dados são cruciais, como em análises financeiras, sistemas de saúde e projetos de data warehousing com grandes volumes de dados.

# **5 Implementação de um exemplo de Modelo Star Schema**

Um modelo Star Schema é uma técnica popular de modelagem de dados em data warehousing e business intelligence. Ele consiste em uma tabela central de fatos, que contém métricas e medidas de negócios, ligada a várias tabelas de dimensão que descrevem os elementos relacionados aos dados. A seguir, como exemplo, vamos utilizar como base um dataset sobre *Segmentação de clientes no Ecommerce*, que pode ser encontrado na plataforma Kaggle, é possível criar um modelo Star Schema como o que está representado abaixo. Após a escolha do dataset, foi feita a identificação da tabela fato, das tabelas dimensões e seus respectivos atributos. Em seguida, foi realizada a construção dos modelos lógico e físico com as devidas relações entre as tabelas.

## **5.1 Modelo Lógico**

****

## **5.2 Modelo Físico**

-- Criação Database

CREATE DATABASE VENDA;

-- Criação table CLIENTE

USE VENDA;

CREATE TABLE CLIENTE(

CLIENTE\_ID INT PRIMARY KEY,

CLIENTE\_PAIS VARCHAR(100) NOT NULL

);

-- Criação table PEDIDO

USE VENDA;

CREATE TABLE PEDIDO(

PEDIDO\_ID INT PRIMARY KEY,

PEDIDO\_QTD INT NOT NULL,

PEDIDO\_VALOR FLOAT NOT NULL

);

-- Criação table FATURA

USE VENDA;

CREATE TABLE FATURA(

FATURA\_NUMERO VARCHAR(20) PRIMARY KEY,

FATURA\_DATA DATE,

FATURA\_HORA TIME

);

-- Criação table TEMPO

CREATE TABLE TEMPO(

TEMPO\_ID INT PRIMARY KEY,

DATA DATE NOT NULL,

HORA TIME NOT NULL

);

-- Criação table FATO\_VENDA

CREATE TABLE FATO\_VENDA(

VENDA\_ID INT PRIMARY KEY,

FK\_CLIENTE INT NOT NULL,

FK\_PEDIDO INT NOT NULL,

FK\_FATURA\_NUMERO VARCHAR(20) NOT NULL,

FK\_TEMPO INT NOT NULL,

FOREIGN KEY (FK\_CLIENTE) REFERENCES CLIENTE(CLIENTE\_ID),

FOREIGN KEY (FK\_PEDIDO) REFERENCES PEDIDO(PEDIDO\_ID),

FOREIGN KEY (FK\_FATURA\_NUMERO) REFERENCES FATURA(FATURA\_NUMERO),

FOREIGN KEY (FK\_TEMPO) REFERENCES TEMPO(TEMPO\_ID)

);

# **6 Cubos OLAP**

Os cubos OLAP (Online Analytical Processing) são estruturas multidimensionais de dados que permitem análise rápida e flexível de informações. Também são conhecidos como cubos multidimensionais ou hipercubos, permitem a visualização e análise de informações de forma multidimensional, o que facilita a compreensão e a tomada de decisões.

Em sua essência, um cubo OLAP é uma matriz multidimensional que contém medidas numéricas (como vendas, lucros, despesas, etc.) e dimensões que categorizam essas medidas (como tempo, produto, região, etc.). Cada célula do cubo representa uma interseção entre uma ou mais dimensões e contém o valor da medida correspondente.

## **6.1 Arquitetura**

Esses cubos são construídos a partir de dados armazenados em bancos de dados OLAP, otimizados para análises complexas e rápidas. Com a ajuda dos cubos OLAP, os usuários podem realizar diversas operações de análise, como drill-down (detalhamento), roll-up (sumarização), slice (fatiamento), dice (segmentação) e pivot (reorganização), para explorar os dados de diferentes perspectivas e responder a perguntas específicas de negócios. Principais componentes de sua arquitetura:

Sua arquitetura geralmente é composta por três camadas principais: a camada de dados, a camada de consulta e a camada de apresentação. Vamos detalhar cada uma:

Camada de Dados é onde os dados são armazenados em um formato multidimensional, geralmente em um banco de dados OLAP. Os dados são organizados em cubos multidimensionais, onde as medidas são representadas em função das dimensões. Os dados são pré-agregados e otimizados para consultas analíticas rápidas, o que diferencia os bancos de dados OLAP dos bancos de dados relacionais tradicionais.

Já a Camada de Consulta será responsável por traduzir as consultas dos usuários em operações no cubo OLAP. Os usuários podem realizar diversas operações analíticas, como drill-down, roll-up, slice, dice e pivot, para explorar os dados de diferentes maneiras. Os sistemas OLAP geralmente incluem um mecanismo de consulta que interpreta as solicitações dos usuários e traduz essas solicitações em consultas no cubo de dados.

Por fim, a Camada de Apresentação será onde os resultados das consultas são apresentados aos usuários de forma compreensível e interativa. Pode incluir interfaces de usuário gráficas, como dashboards, relatórios, gráficos e tabelas dinâmicas, que permitem aos usuários visualizar e interagir com os dados analisados.

Essas camadas trabalham juntas para fornecer aos usuários uma experiência de análise de dados eficiente e intuitiva. Os cubos OLAP e sua arquitetura são projetados para facilitar a análise multidimensional de grandes volumes de dados e ajudar as organizações a tomar decisões informadas com base em insights extraídos desses dados.

## **6.2 Implementação**

A implementação dos cubos OLAP envolve várias etapas, desde a modelagem dos dados até a construção e a disponibilização para consulta. A seguir vamos ter uma visão geral do processo:

* Modelagem Dimensional:

O primeiro passo é projetar o modelo dimensional dos dados. Isso envolve identificar as dimensões que serão utilizadas para categorizar os dados (como tempo, produto, cliente, etc.) e as medidas que serão analisadas (como vendas, lucro, quantidade, etc.).

O modelo dimensional geralmente é representado por meio de esquemas estrela ou floco de neve, nos quais uma tabela de fatos (contendo as medidas) está conectada a várias tabelas de dimensão (contendo os atributos das dimensões).

* Extração, Transformação e Carregamento (ETL):

Os dados são extraídos das fontes de dados operacionais, como bancos de dados transacionais, arquivos CSV, planilhas, etc.

Em seguida, os dados são transformados para se adequarem ao modelo dimensional e para garantir consistência, integridade e qualidade.

Por fim, os dados transformados são carregados no banco de dados OLAP, onde serão armazenados e utilizados para construir os cubos.

* Construção dos Cubos OLAP:

Com os dados carregados no banco de dados OLAP, os cubos OLAP são construídos. Isso geralmente é feito por meio de ferramentas específicas de construção de cubos fornecidas pelos fornecedores de software OLAP.

Durante a construção dos cubos, são definidas as hierarquias das dimensões, as agregações necessárias para otimizar o desempenho das consultas e outras configurações relevantes.

* Disponibilização para Consulta:

Uma vez que os cubos são construídos, eles estão prontos para serem consultados pelos usuários.  
As consultas aos cubos OLAP podem ser realizadas por meio de interfaces de consulta específicas (como ferramentas de BI ou linguagens de consulta OLAP) que permitem aos usuários explorar os dados de forma interativa e analisar informações de diferentes perspectivas.

* Manutenção e Atualização:

Os cubos OLAP requerem manutenção regular para garantir que os dados estejam atualizados e que o desempenho das consultas seja otimizado.

Isso pode envolver a atualização dos dados de origem, a reconstrução dos cubos com novos dados e ajustes nas configurações de agregação e indexação conforme necessário.

## **6.3 Casos de uso**

### **6.3.1 Análise de Vendas e Desempenho Financeiro**

Os cubos OLAP são frequentemente empregados para análise de vendas, permitindo aos usuários explorar dados de vendas por região, produto, canal de vendas, período de tempo, entre outras dimensões. Eles também são úteis para análise financeira, como monitoramento de lucros, custos, orçamentos e previsão de receitas.

### **6.3.2 Gestão de Estoque e Cadeia de Suprimentos**

Na área de gestão de estoque e cadeia de suprimentos, os cubos OLAP ajudam na análise de inventário, demanda por produtos, ciclo de vida do produto, tempo de entrega e eficiência da cadeia de suprimentos. Isso permite uma gestão mais eficaz do estoque e uma resposta mais ágil às demandas do mercado.

### **6.3.3 Análise de Clientes e Segmentação de Mercado**

Os cubos OLAP são úteis para análise de clientes, permitindo a segmentação de clientes com base em diferentes características (como comportamento de compra, perfil demográfico, preferências, etc.). Isso ajuda as empresas a personalizar suas estratégias de marketing, melhorar a fidelidade do cliente e identificar oportunidades de crescimento de mercado.

### **6.3.4 Planejamento e Orçamento Empresarial**

Na área de planejamento e orçamento empresarial, os cubos OLAP são usados para análise de orçamento, previsão financeira, acompanhamento de desempenho em relação às metas e simulação de cenários. Isso permite uma gestão mais eficiente dos recursos financeiros e uma tomada de decisão mais informada.

### **6.3.5 Análise de Recursos Humanos**

Os cubos OLAP são aplicados na análise de recursos humanos para monitorar métricas como rotatividade de funcionários, satisfação dos funcionários, eficácia de programas de treinamento e desenvolvimento, entre outros. Isso ajuda na gestão estratégica de recursos humanos e no desenvolvimento de políticas e práticas de RH mais eficazes.

### **6.3.6 Análise de Dados de Saúde**

Na área da saúde, os cubos OLAP são utilizados para análise de dados clínicos, gestão de pacientes, monitoramento de indicadores de saúde pública, análise de eficácia de tratamentos, entre outros. Isso contribui para uma melhor gestão de recursos e melhoria dos cuidados de saúde.

# **7 Implementação de um Exemplo de Cubo OLAP no MySQL**

A partir do Dataset sobre Vendas em um E-commerce - que foi explorado neste estudo - , em formato .CSV, e inserido no MySQL Workbench, é possível realizar DQLs que permitem ao MySQL simular um Cubo OLAP. Vejamos:

## **7.1 Consulta**

-- Criar uma tabela virtual representando o cubo OLAP

CREATE VIEW Cubo\_Faturas AS

SELECT

Numero\_da\_fatura,

ID\_Cliente,

Pais,

Data\_da\_fatura,

Hora\_da\_fatura,

SUM(Quantidade) AS Total\_Quantidade,

SUM(Valor) AS Total\_Valor

FROM

Faturas

GROUP BY

Numero\_da\_fatura,

ID\_Cliente,

Pais,

Data\_da\_fatura,

Hora\_da\_fatura;

-- Consulta para obter a soma total de valor por país e data

SELECT

Pais,

Data\_da\_fatura,

SUM(Total\_Valor) AS Total\_Valor\_Por\_Pais\_Data

FROM

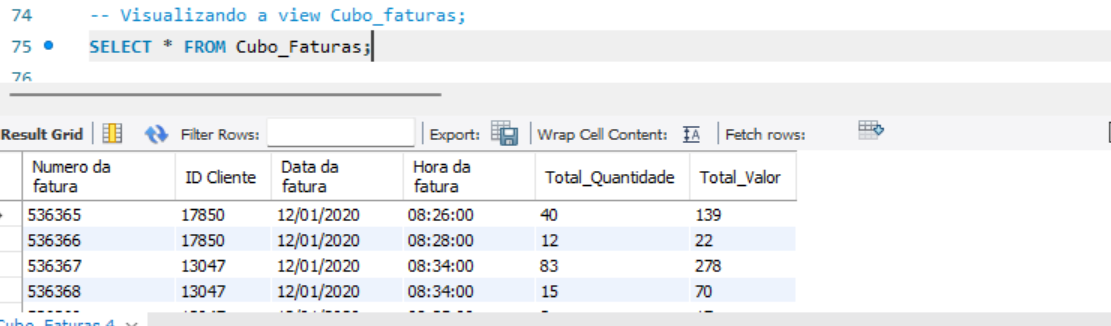
Cubo\_Faturas

GROUP BY

Pais,

Data\_da\_fatura;

## **7.2 Resultado**

****

## **7.3 Consulta**

-- Consulta para obter a soma total de valor por cliente e data

SELECT

`ID Cliente`,

`Data da fatura`,

SUM(Total\_Valor) AS Total\_Valor\_Por\_Cliente\_Data

FROM

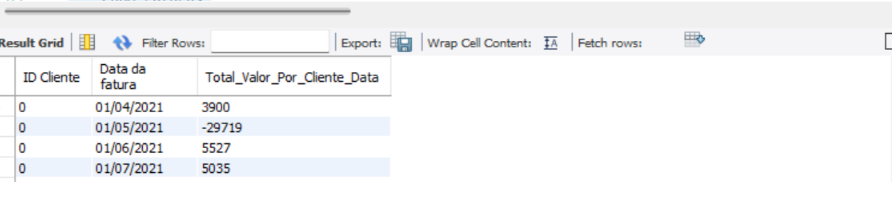
Cubo\_Faturas

GROUP BY

`ID Cliente`,

`Data da fatura`;

## **7.4 Resultado**

****

# **Referências Bibliográficas**

ELMASRI, R., NAVATHE, S. B. Sistemas de Banco de Dados. Visão geral de data warehousing e OLAP. s.l, p. 995-1006, 7ª Edição. SP. Ed. Pearson, 2018.

NASCIMENTO, Filipe. Database Application. Aula 06. 21 mai. 2024. Apresentação do Power Point. Disponível em: 3º Período - TI - Turma A T.I Noite 2024.1. Acesso em 21 mai. 2024.