

# Super Store

ROTA: Estrutura de dados - ETL

Socorro Moura



◁ Laboratorio ▷

# Contextualização do Projeto

A Super Store, como líder no varejo, lida com um volume massivo de dados dispersos e complexos.

Para transformar essa realidade e otimizar a tomada de decisões, implementamos um robusto sistema ETL (Extract, Transform, Load).

Este sistema organiza nossos dados em tabelas de fatos e dimensões, garantindo análises mais rápidas e insights acionáveis.



Meta

# Objetivo

**Nosso objetivo foi criar um sistema ETL (Extract, Transform, Load) robusto para a Super Store.**

**Focamos em extrair dados de diversas fontes, transformá-los conforme nossas necessidades e carregá-los em um data warehouse organizado.**

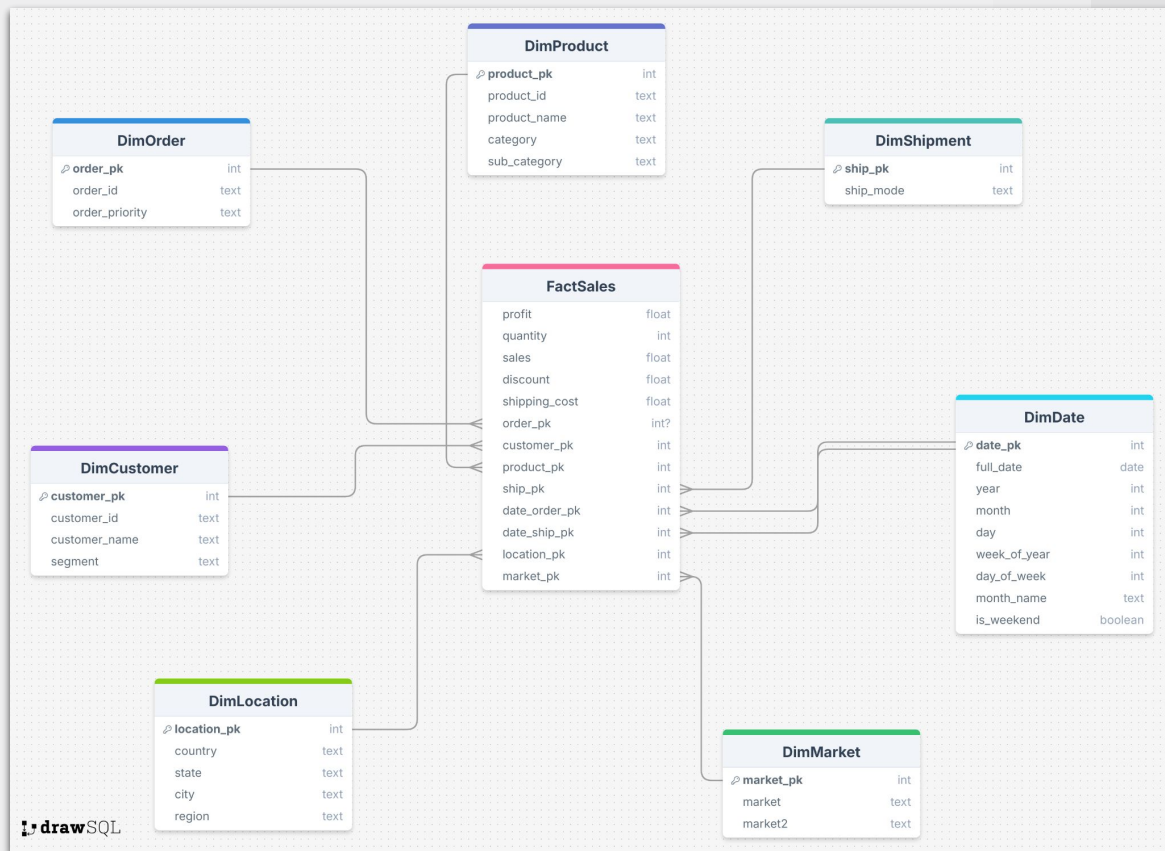
ETL

# Estrutura

Sistema ETL:

Tabelas de fatos e dimensões

O modelo escolhido foi o **estrela** (Star Schema) que consiste em uma tabela de fatos central e várias tabelas de dimensão que a cercam.

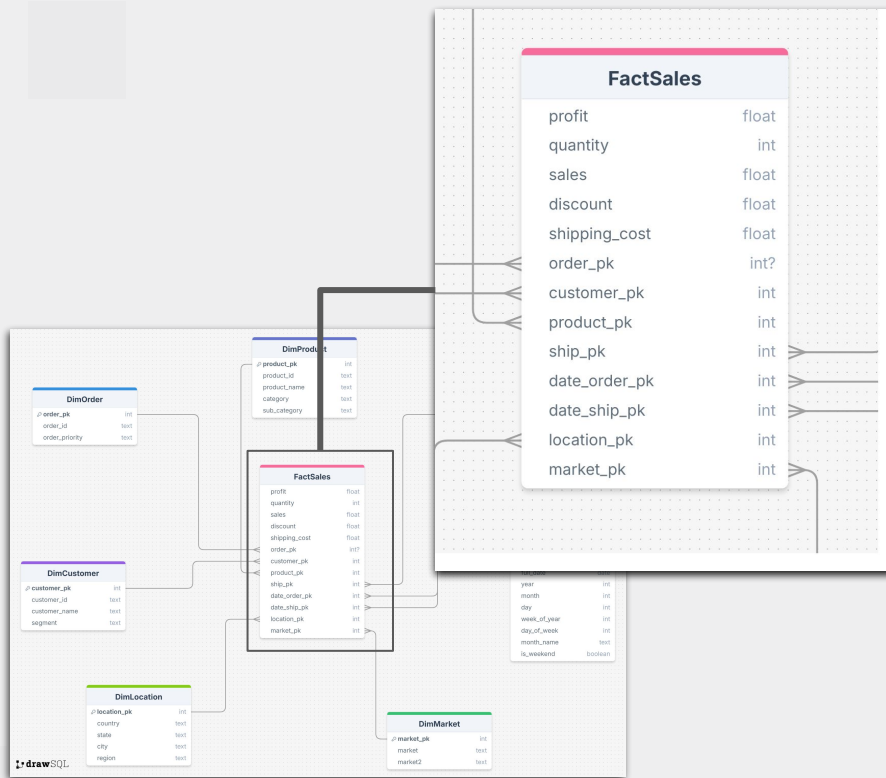


## ESTRUTURA

# Tabela de Fatos

Contém as métricas quantitativas (os "fatos") do seu negócio e chaves estrangeiras para as tabelas de dimensão

- Tabela central contém as **métricas de vendas** (o evento mais importante) e as chaves que ligam às dimensões;
- O que ela guarda? Coisas que você pode **contar** ou **somar**:
  - Quantos produtos foram vendidos? (**quantity**)
  - Quanto dinheiro entrou? (**sales**)
  - Qual foi o lucro? (**profit**)
  - Qual desconto foi dado? (**discount**)
  - Qual foi o custo de envio? (**shipping\_cost**)
- Também contém Chaves Estrangeiras (FKs) que a ligam às outras tabelas;

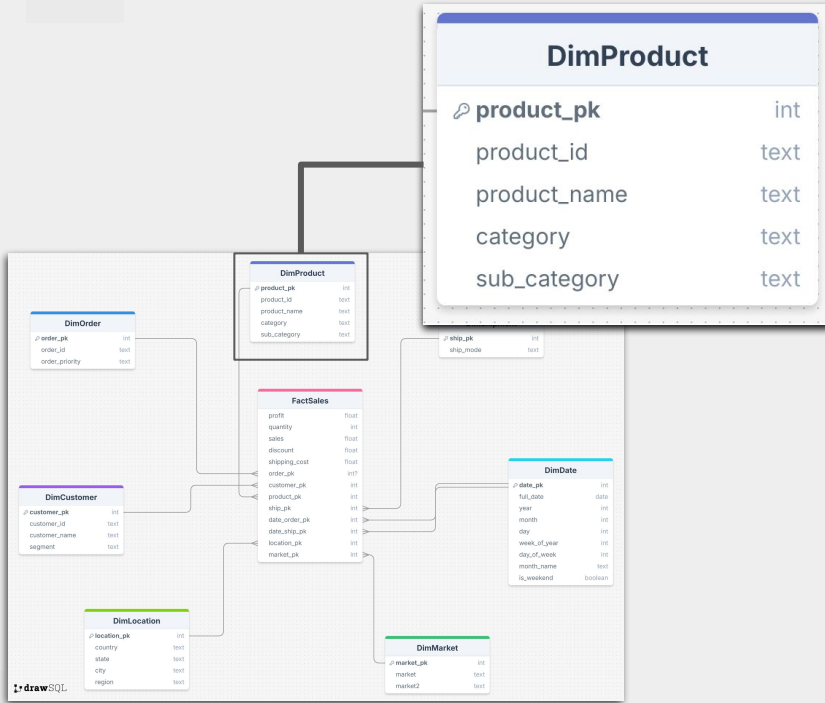


## ESTRUTURA

# Tabelas de Dimensão

Contêm os atributos descritivos sobre as entidades do seu negócio (quem, o que, onde, quando, como). Elas fornecem o "contexto" para os seus fatos.

- Guardam **detalhes** sobre as coisas
- Respondem perguntas como:
  - **QUEM** comprou? (Clientes)
  - **O QUÊ** foi comprado? (Produtos)
  - **ONDE** aconteceu? (Cidades, Países)
  - **QUANDO** aconteceu? (Datas)
  - **COMO** foi enviado? (Modo de Envio)
- Cada item aqui tem um **código único** (a Chave Substituta - PK). É como o código de barras do seu produto que garante que ele é único naquela prateleira.



AGILIDADE

# Por que tudo isso?

Essa organização é como arrumar a sua loja de brinquedos em prateleiras lógicas. Isso faz com que seja **muito mais fácil e rápido** responder a perguntas de negócios complexas, como:

- "Qual o lucro total da categoria 'Móveis' na 'América do Norte' no ano passado?"
- "Quantos produtos foram enviados com 'Entrega Expressa' para a cidade de 'São Paulo' em Março?"

Sem essa organização, você teria que vasculhar toda a sua "loja bagunçada" para cada pergunta.

Com ela, você **vai direto** nas prateleiras certas, pega os **números** da *FactSales* e os **detalhes** das *Dimensões* correspondentes.

SUPER STORE

# Processo de ETL

**Pipeline de Atualização do Data Warehouse**



CAMINHOS

# O Processo ETL: da Origem à Análise

01



## 1. Preparação dos Dados

- Extração da fonte original.
- Transformação para limpeza e padronização (ex: texto em maiúsculas).
- Criação de uma base limpa (superstore\_standardized).

02



## 2. Organização: Carregamento das Dimensões

- Dados únicos carregados em tabelas de dimensão (clientes, \* produtos, datas, etc.).
- Criação de códigos únicos (PKs) para cada detalhe.
- Isso nos dá o contexto para análises.

03



## 3. Conexão: Carregamento da Tabela de Fatos (FactSales)

- Preenchimento da FactSales com métricas de vendas (quantidade, lucro, etc.).
- Conexão com as dimensões através de chaves estrangeiras (FKs).
- Resulta em um Data Warehouse otimizado para insights estratégicos.

ATENÇÃO

# Recomendações

## Automação do ETL:

Implementar ferramentas para automatizar o fluxo de dados, garantindo atualizações consistentes e menos erros.

## Cargas Incrementais:

Processar apenas dados novos ou alterados, otimizando performance e custos.

## Dashboards e Relatórios:

Conectar o DW a ferramentas de BI para criar painéis visuais e gerar insights rapidamente.

## Expansão de Fontes:

Integrar novos dados (ex: marketing) para uma visão 360° do negócio.

INSIGHTS

# Conclusões

Este projeto entregou um Data Warehouse **robusto** e **otimizado** no BigQuery.

Com um ETL eficiente, transformamos dados dispersos em uma base confiável e estruturada.

Agora, a Super Store tem a agilidade e os insights necessários para decisões estratégicas e para manter sua **vantagem competitiva** no mercado.



ATÉ A PRÓXIMA

# Obrigada pela atenção!

[mariadosocorro.m@gmail.com](mailto:mariadosocorro.m@gmail.com)

<https://github.com/socorromouras>

<https://www.linkedin.com/in/mariadosocorrms>