

Paradigmas OLAP e OLTP

Grupo de Estudos de
Engenharia de Dados

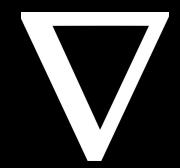
 /caue-paiva
Cauê Paiva Lira

Significado das siglas

- **Online Transaction Processing (OLTP)**
- **Online Analytical Processing (OLAP)**

Online Processing: Ambos paradigmas vieram a tona com a explosão de dados e usuários da internet, especialmente nos anos 90 em diante.

Analytical x Transactional: O principal diferencial entre os paradigmas é o caso de uso/regra de negócio que se pretende suportar.



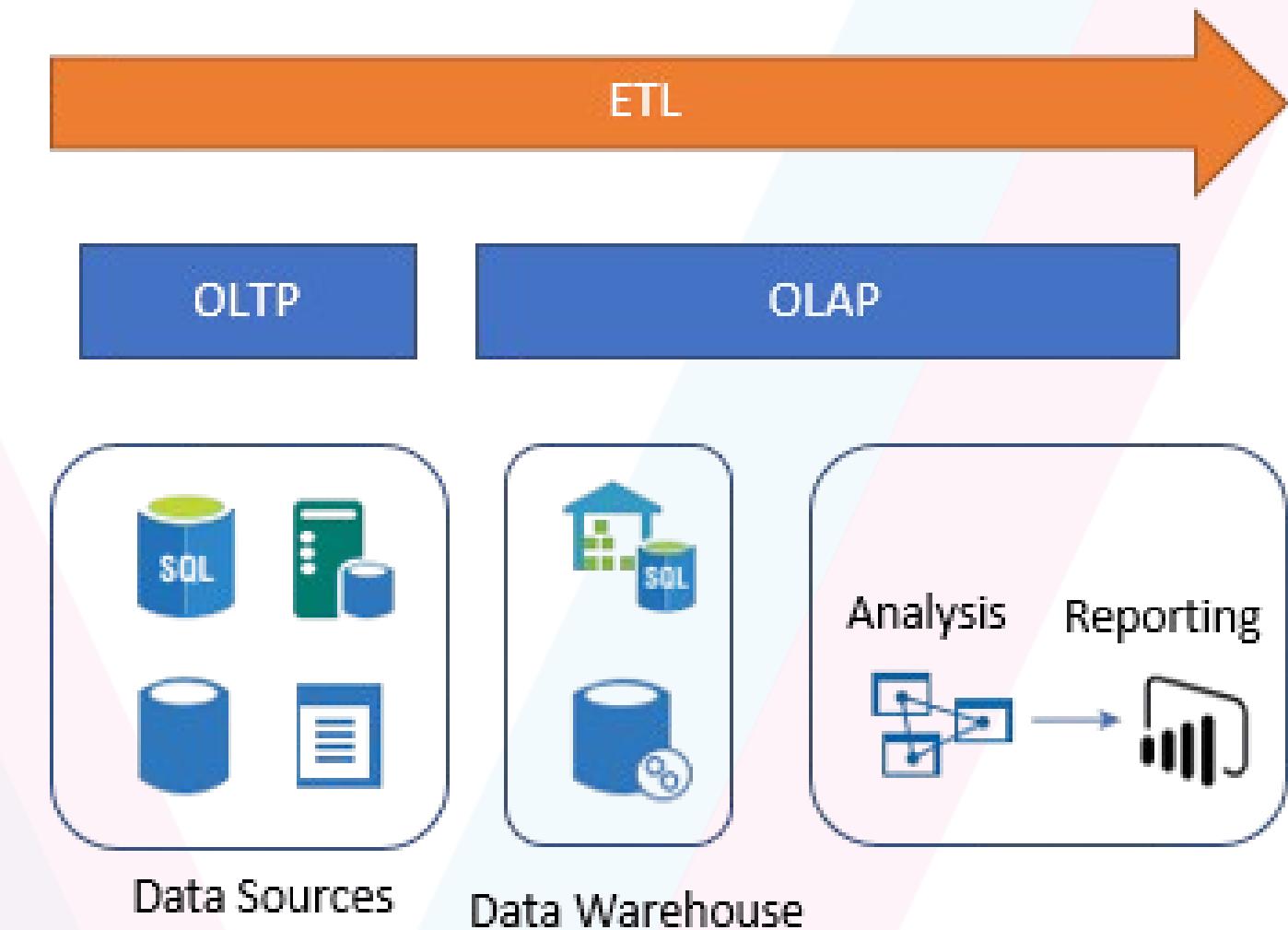
Analytical x Transactional

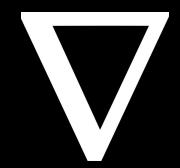
Transactional: Foco em **transações simples, frequentes e consistentes**, geralmente apoioando regras cruciais do negócio. Os dados são **atuais**.

Ex: Criar ou alterar uma conta, colocar um item no carrinho, realizar um compra online.

Analytical: Foca em **análise de dados históricos**, com transações mais complexas porém menos frequentes. Menor necessidade de consistência.

Ex: Criar Dashboards para analistas, calcular média de vendas em um período.

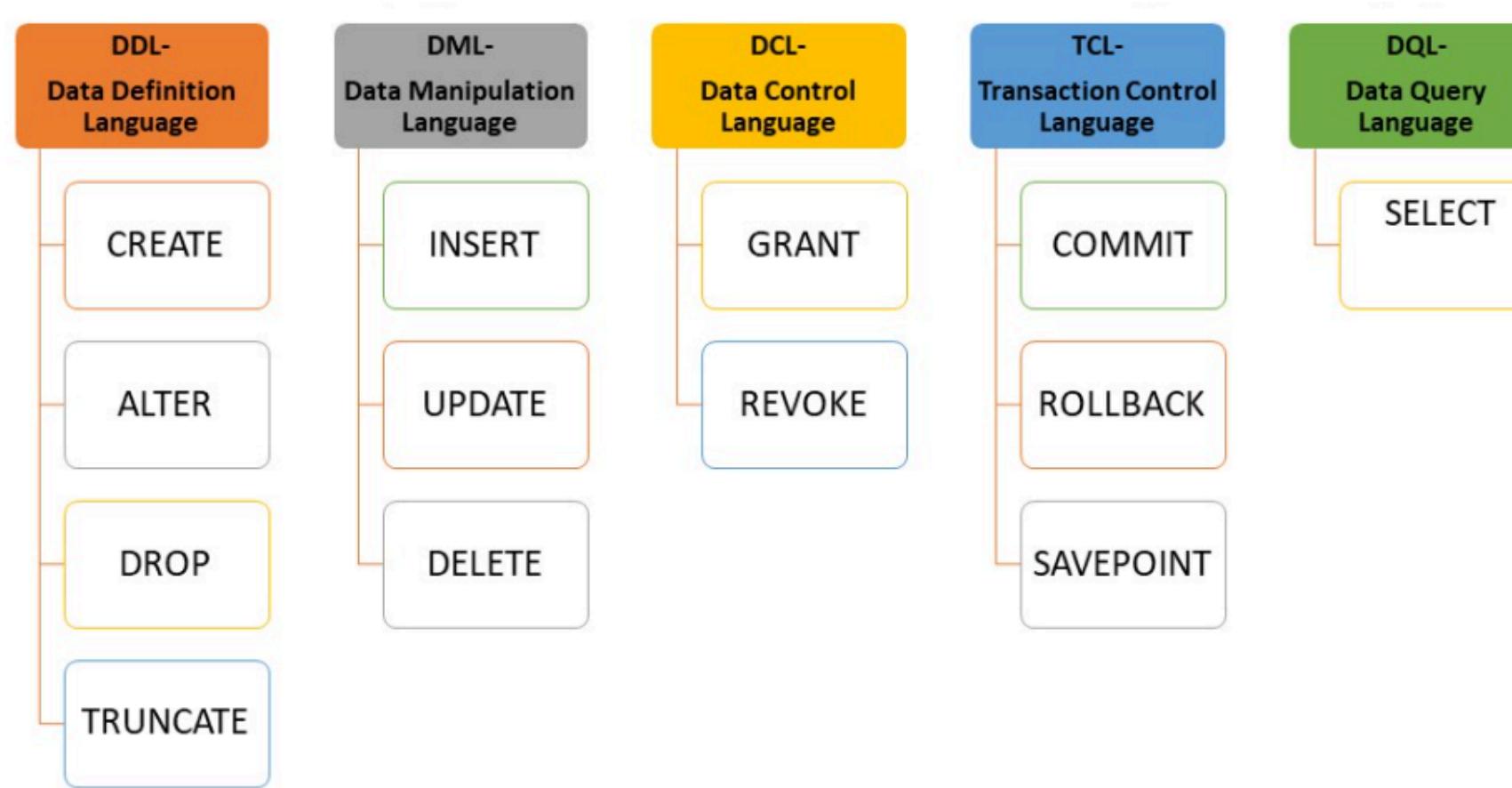


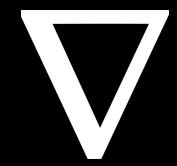


SQL

A linguagem SQL é comumente utilizada na implementação de **sistemas OLTP e OLAP**. Não somente em BD relacionais mas também em outras ferramentas (Ex: Spark SQL)

Uma forma de ilustrar a **diferença entre ambos paradigmas** é analisar certas operações da linguagem SQL



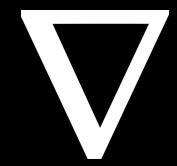


DML e DQL

DML (Data Manipulation Language): Comandos do SQL que lidam com manipulação de dados. Ex: **INSERT, UPDATE e DELETE**

DQL (Data Query Language): Comando do SQL para fazer queries/buscas em tabelas. Ex: **SELECT**

Com base nesse subconjunto de **comandos**, vamos analizar como sua **utilização difere nos sistemas OLAP e OLTP**



Diferença nos comandos SQL

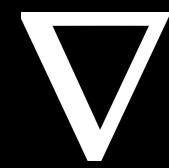
OLTP, Comandos mais comuns:

- INSERT (Dados individuais)
- UPDATE
- DELETE
- SELECT (Valores específicos)

OLAP, Comandos mais comuns:

- INSERT (Dados em Batch por meio de um processo de ETL)
- SELECT (Agregações, Contas...)

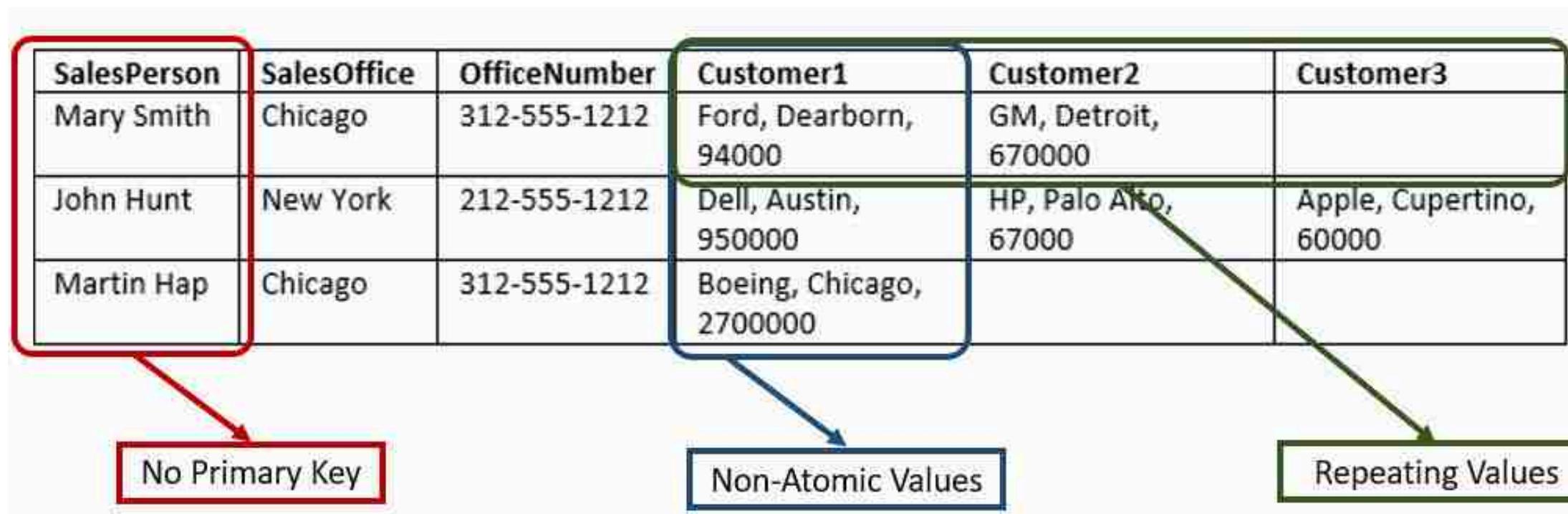
Sistemas **OLAP** geralmente **não precisam se preocupar em atualizar ou deletar registros**, pois lidam com dados históricos já processados



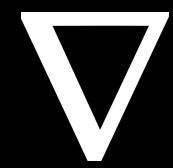
Modelagem OLTP

Banco de Dados **OLTP** seguem uma modelagem baseado no modelo **entidade-relacionamento**, geralmente **normalizado** na 3NF (Terceira Forma Normal)

Principais Objetivos da modelagem: Evitar duplicação de dados, garantir consistência durante Inserções, Updates e Deletes



-Exemplo dos problemas de uma tabela não normalizada

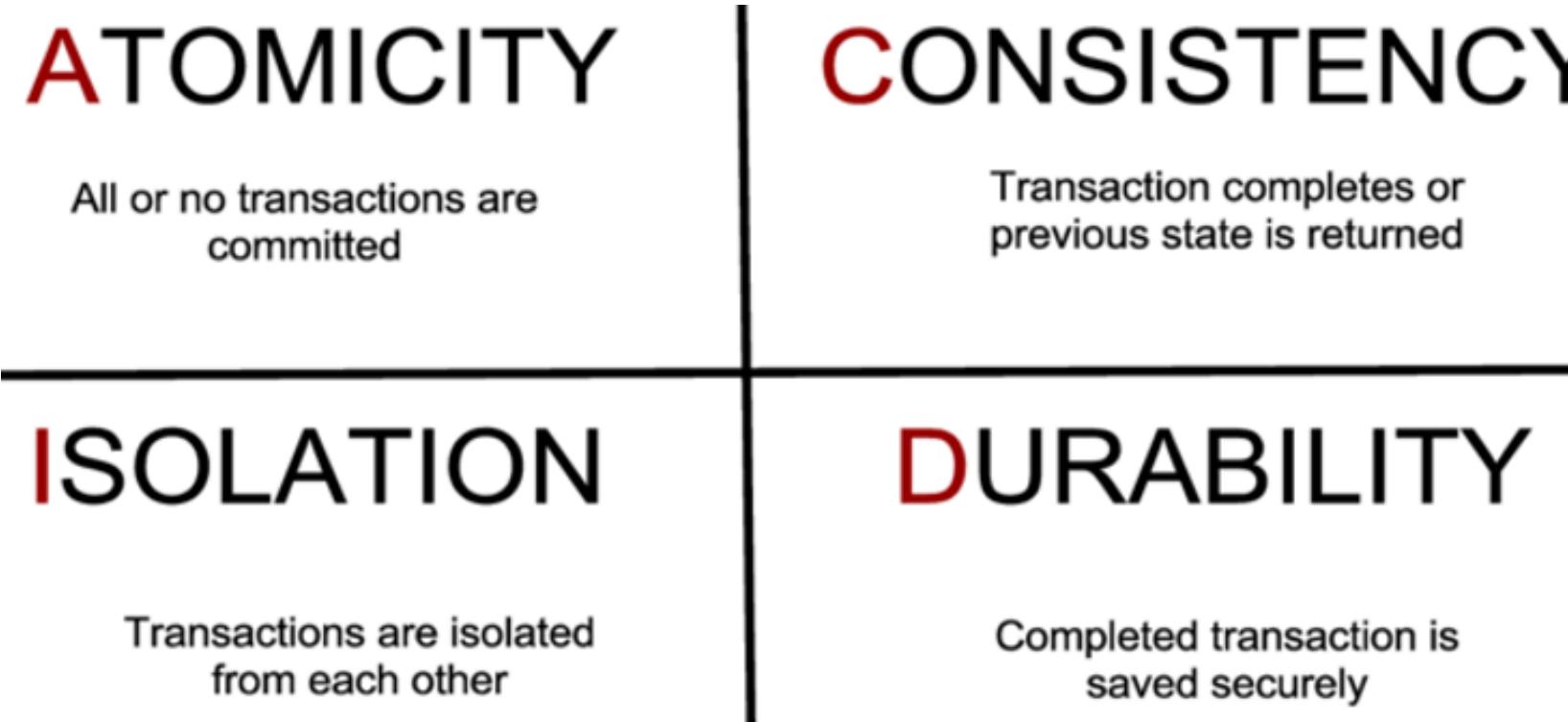


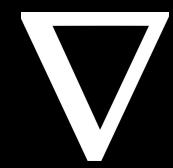
Tecnologias para sistemas OLTP

Sistemas OLTP são geralmente criados a partir de **SGBDs relacionais** (Postgres, MySQL, SQL server...)



Os BDs desse sistema devem respeitar os **princípios ACID**





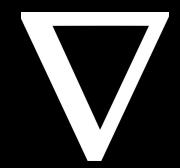
Modelagem OLAP

Sistemas OLAP seguem uma **modelagem dimensional**, também conhecida como esquema estrela.

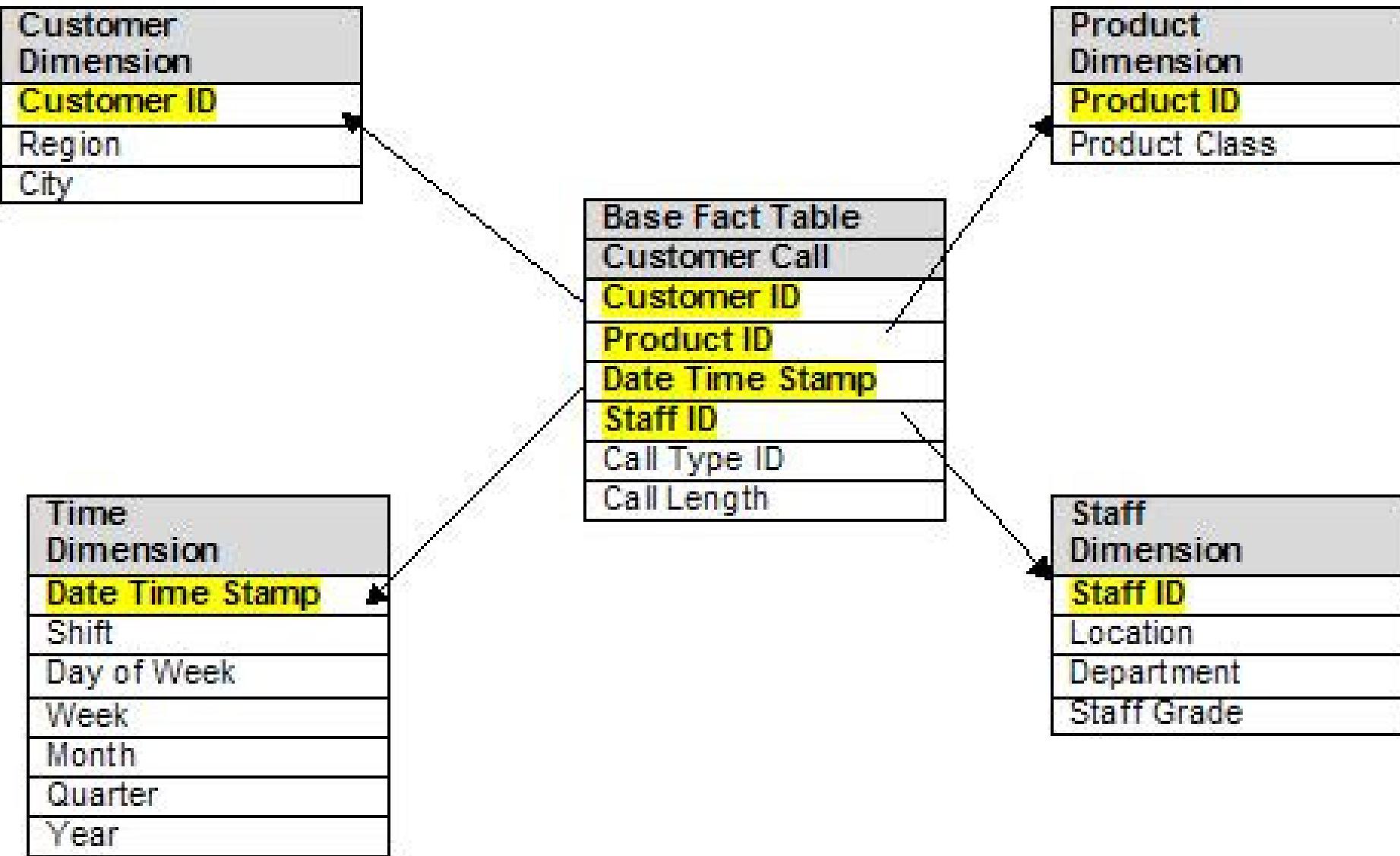
Nesse Schema o BD é divido em 2 tipos de tabelas:

- **Tabelas Dimensão:** Guardam Atributos usados em queries por analistas para summarização/Agregação. Ex: Data, Local...
- **Tabelas de Fatos:** Guardam os dados históricos que serão analizados. Esses dados geralmente vem de alguma atividade do negócio

Características: Desnormalização, Menos Joins entre tabelas, apoio a Queries mais complexas

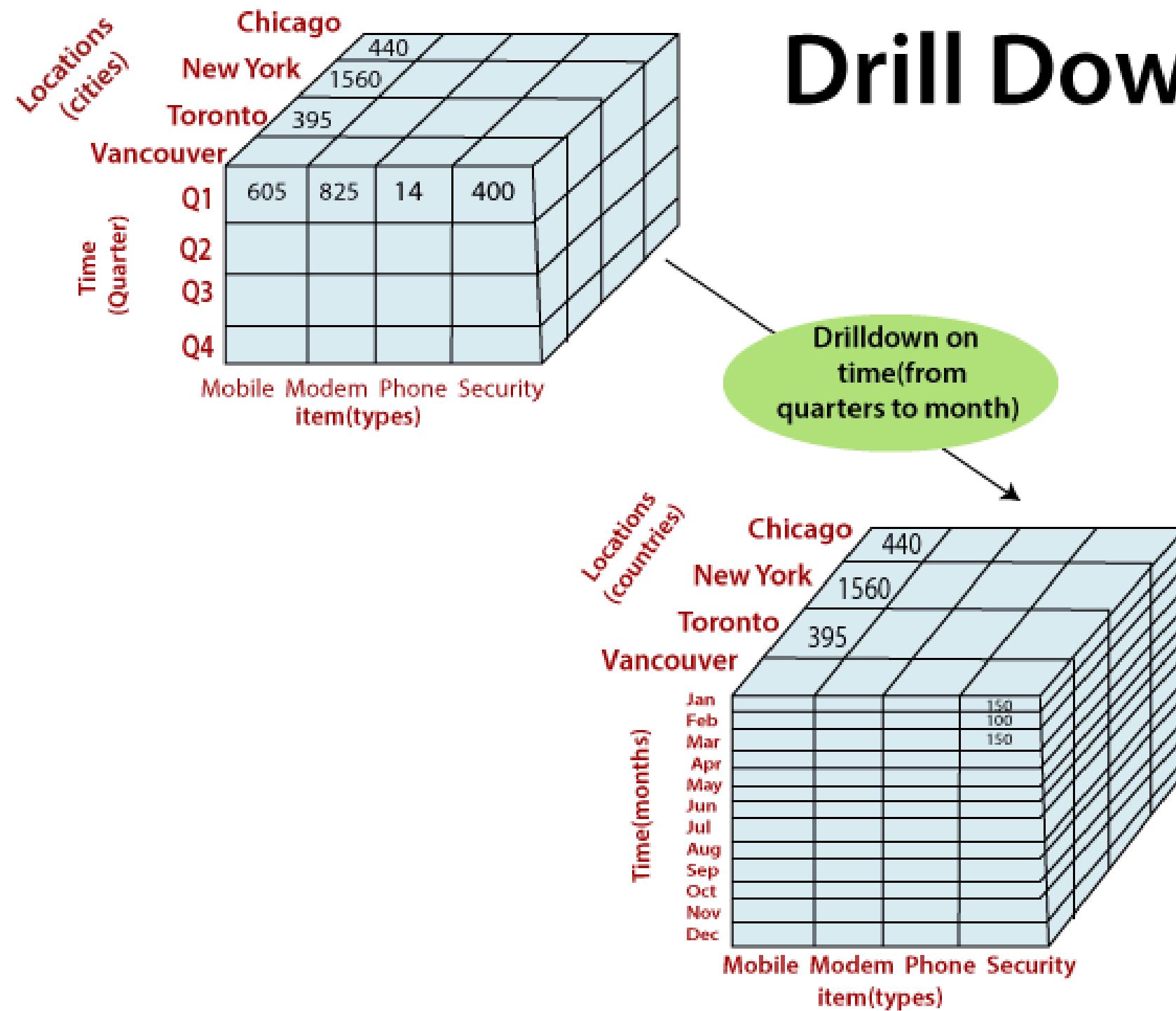


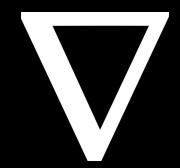
Star Schema



-Exemplo de uma modelagem no padrão estrela

Drill Down





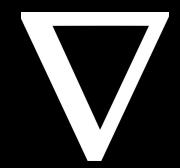
Tecnologias para sistemas OLAP

Sistemas OLAP **podem ser implementados com BDs relacionais**,
porém seu caso de uso específico e o fato de não precisar conformar
com os princípios ACID levam à **outras tecnologias sendo utilizadas**

BDs orientados a **Colunas**:

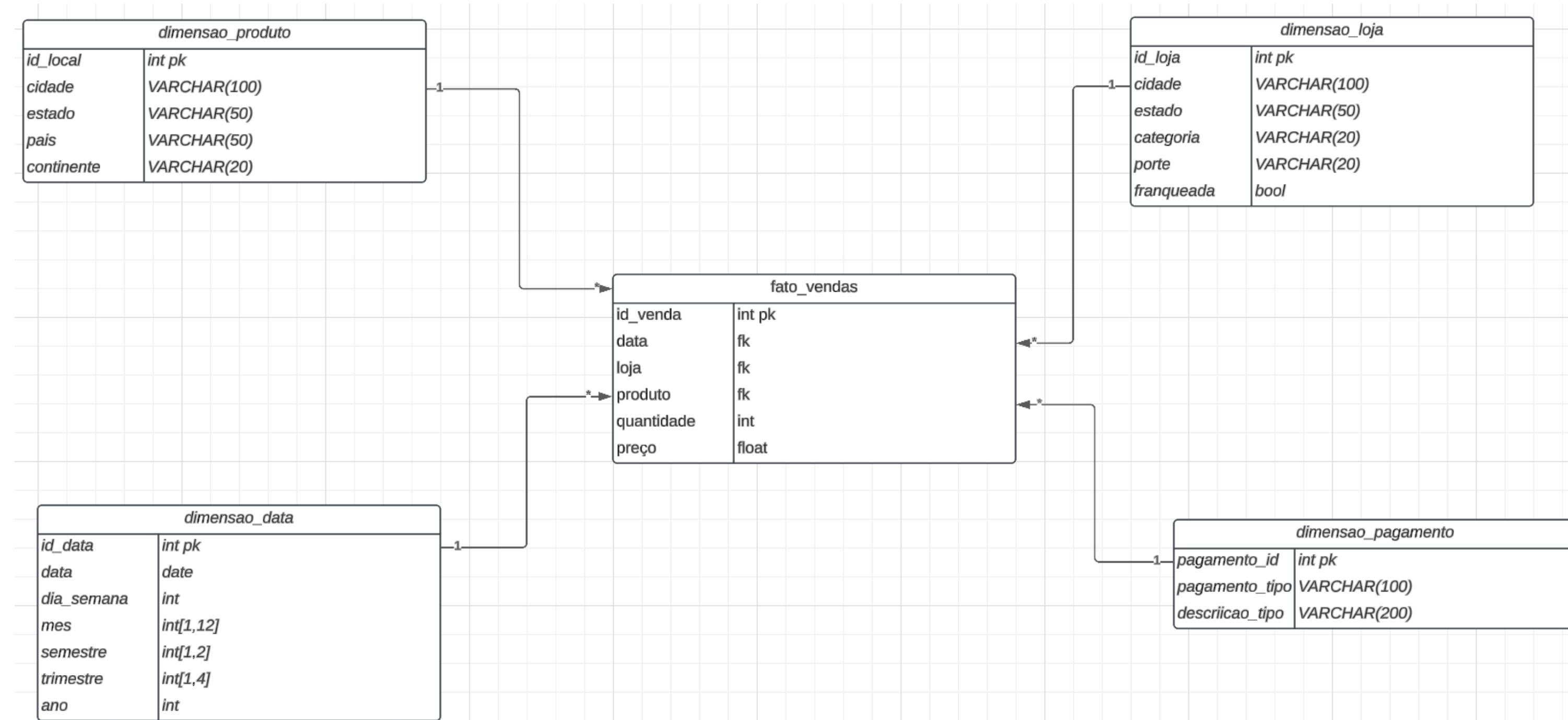


Amazon Redshift

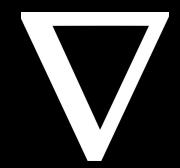


Mão na Massa

Vamos criar um schema de um Data Warehouse no PostgreSQL, inserir dados sintéticos nele e fazer algumas queries analíticas.



-Schema que vamos construir

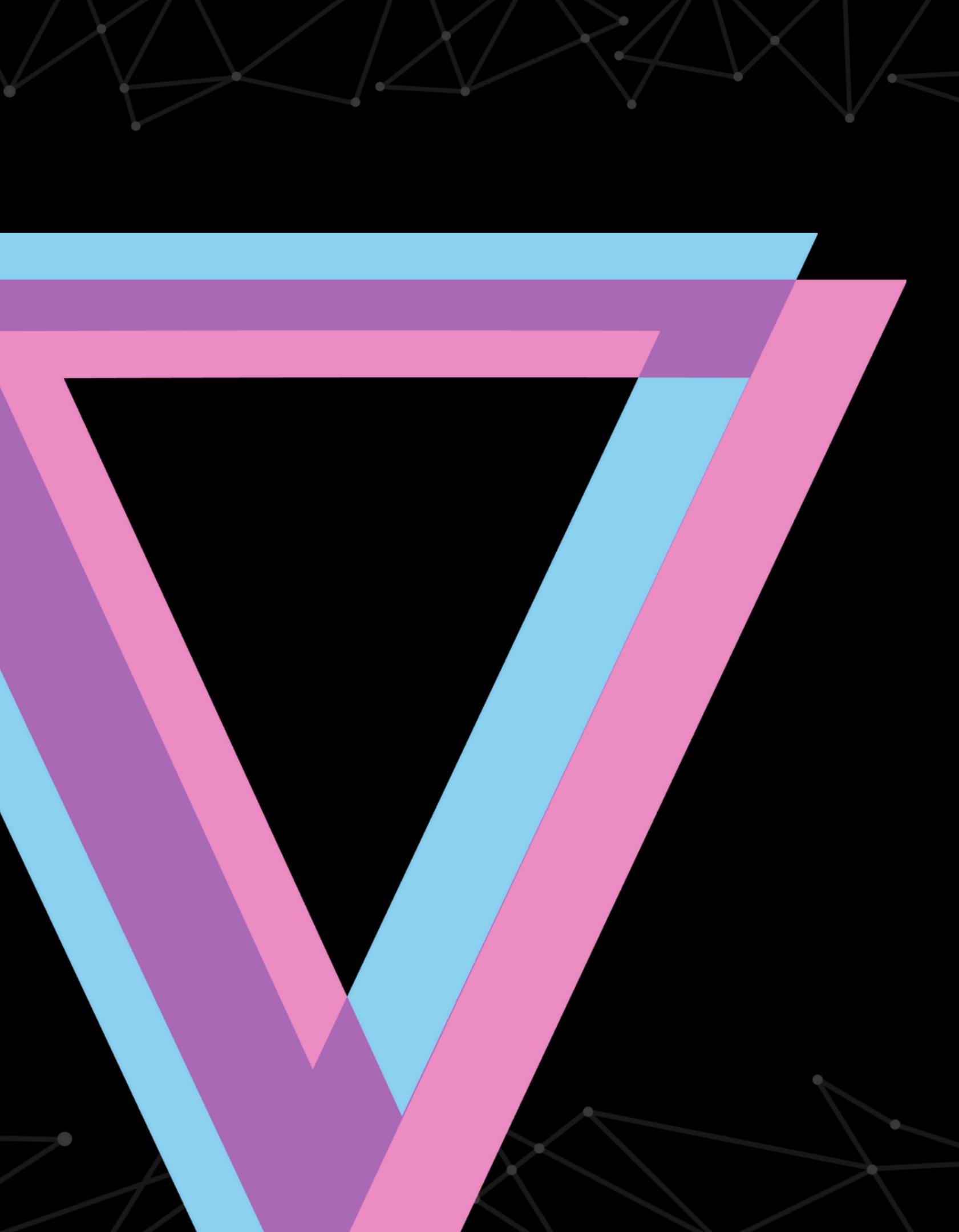


Mão na Massa

Código da demonstração disponível no Github



<https://github.com/caue-paiva/data warehousing studies>



Obrigado pela atenção!

Grupo de Estudos de
Engenharia de Dados

 /caue-paiva
Cauê Paiva Lira