

# Arquitetura de Dados Analíticos

Apresentação da Disciplina



# Professor

Gustavo Dias

Mestre em Gestão e Desenvolvimento Regional

Consultor em Business Intelligence, com atuação em projetos na área da saúde e varejo.

## Contatos:

 p.gustavodias@gmail.com

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

 pgustavodias



# Agenda de aula

- Ementa;
- Objetivos;
- Distribuição de notas;
- Acesso ao conteúdo;
- Introdução

Conhecimentos desenvolvidos





# Ementa

- Conceitos de Data Warehouse e Data Lake;
- Modelagem dimensional e relacional;
- Esquemas Estrela, Floco de Neve e Constelação;
- Integração de dados e o processo ETL;
- Dimensionamento e otimização de modelos;
- Gerenciamento de dados estruturados e não estruturados;
- Monitoramento, otimização, segurança e conformidade.

## Projeto Analítico





# Objetivo

Capacitar o estudante a compreender, projetar e implementar ambientes de dados para análise, envolvendo Data Warehouse e Data Lake, com foco em modelagem, integração, armazenamento, segurança e desempenho de dados.

## Projeto Analítico



# Distribuição de Notas

**N1 (30)**

**Avaliação 1**

Modelagem  
de ambientes  
analíticos

**N2 (30)**

**Avaliação 2**

Implementação  
de ambientes  
analíticos

**N3 (20 + 20)**

Projeto de  
Data  
Warehouse e  
Data Lake

# Compartilhamento de Materiais



[pgustavodias](#)

[arquitetura-de-dados-analiticos](#)



.Introdução

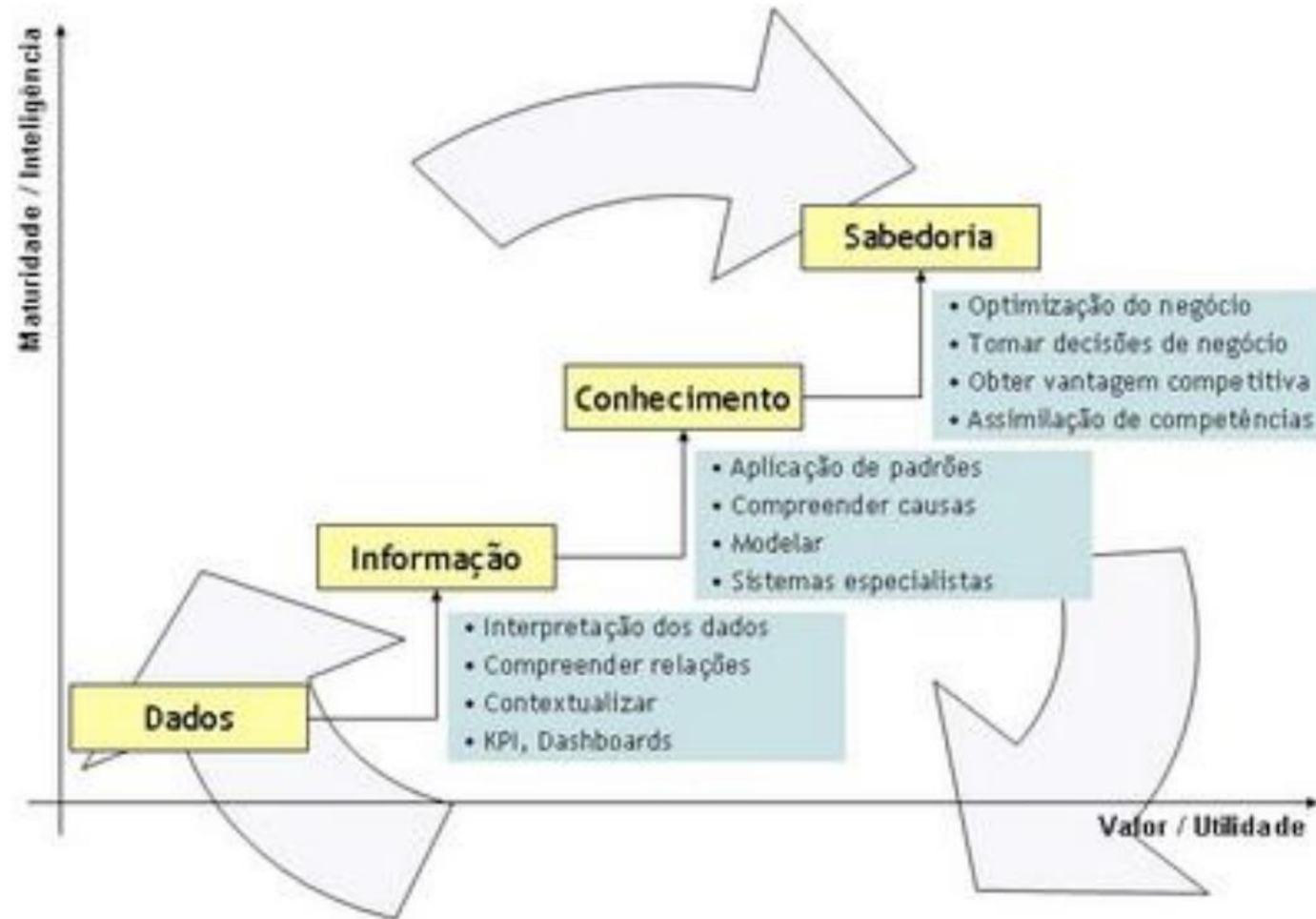


# Introdução

Para você o que  
é um dado?



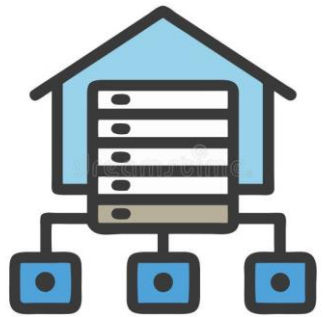
# Introdução



“ Dados são fluxos de fatos brutos que representam eventos ocorridos nas organizações ou no ambiente físico antes de serem organizados e arranjados em uma forma que as pessoas possam compreender e utilizar. ”

LAUDON, K. C.; LAUDON, J. P.

# Data Warehouse, Data Lake e o ETL



DW e DL são formas de armazenar os dados (estruturados e não estruturados) para transformá-los em informação. O processo ETL (ou ELT) viabiliza o enriquecimento e o armazenamento.





# Conceitos Importantes

**Modelo de (Banco de) Dados**: descrição formal da estrutura de um banco de dados.

Para se construir um modelo de dados, usa-se uma **linguagem de modelagem de dados**, podendo ser textual ou gráfica.

Existem diferentes técnicas e conceitos a serem aplicados, que caracterizarão a **abordagem de modelagem**.

Abstrações mais comuns: **modelo conceitual e modelo lógico**.



# Abordagens

Há diferentes abordagens para a construção de modelos de dados. Dentre elas destacam-se:

- a) Top-Down:
  - a) Parte do modelo conceitual (estratégico);
  - b) Alinhada aos objetivos organizacionais;
  - c) Adequado para projetos corporativos (ERPs);
  - d) Reduz redundâncias.

## Reforço do aprendizado



Cliente, pedido e produto surgem primeiro.  
Em seguida detalham-se campos como CPF, data, valores etc.

# Abordagens

Há diferentes abordagens para a construção de modelos de dados. Dentre elas destacam-se:

b) Botton-Up:

- a) Parte de modelos existentes;
- b) Aplicada à projetos de inteligência;
- c) Mais rápido para ambientes em operação;
- d) Risco semântico caso não haja organização.

## Reforço do aprendizado



Analisa-se as tabelas já existentes.  
Posteriormente define-se entidades e suas relações.

# Sistema Gerenciador de Banco de Dados (SGDB)

É um software que incorpora as funções de definição, recuperação e alteração em um banco de dados, promovendo o armazenamento e o acesso aos dados.





# SGDB



## **Modelo OLAP**

Foco em análise, não em transação



## **Armazenamento Colunar**

Otimiza consulta e compressão



## **Compressão Avançada**

Ganhos em leitura e cache = menos disco mais performance



## **Processamento Paralelo Massivo (MPP)**



# SGDB



**Otimização para Agregação**  
Agregação é o padrão



**Menor Normalização**  
Modelagem dimensional



**Gestão de Concorrência**  
Prioriza a leitura e não a atualização imediata



**Carga de Dados em Lote**  
Lote, incrementais ou particionadas



# Síntese Comparativa

## **Característica**

Armazenamento

Leitura

Escrita

Modelo

Escalabilidade

Foco

## **OLTP**

Linha

Pontual

Frequente

Normalizado

Vertical

Operação

## **OLAP**

Coluna

Massiva

Lote

Dimensional

Horizontal

Decisão

# Referências

DATE, C. J. Introdução a sistemas de bancos de dados. 8. ed. Rio de Janeiro: Elsevier, 2004.

HEUSER, Carlos Alberto. Projeto de banco de dados. 6. ed. Porto Alegre: Bookman, 2009.

INMON, William H. Building the Data Warehouse. 4. ed. Hoboken: John Wiley & Sons, 2005.

KIMBALL, Ralph; ROSS, Margy. The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling. 3. ed. Hoboken: John Wiley & Sons, 2013.

LAUDON, Kenneth C.; LAUDON, Jane P. Sistemas de informação gerenciais. 11. ed. São Paulo: Pearson, 2014.

RIBEIRO, Fernanda; SANTOS, Plácida Leopoldina Ventura Amorim da Costa. Dados, informação e conhecimento: conceitos e relações. São Paulo: Cultura Acadêmica, 2020.

TORINO, Emanuelle. Arquitetura de dados no contexto da Ciência da Informação. 2023. Tese (Doutorado em Ciência da Informação) – Universidade Estadual Paulista (UNESP), Marília, 2023.



# Referências

## Aplicação:

- Heuser (2009) → base estrutural da arquitetura de SGBDs, incluindo modelagem conceitual, lógica e física, servindo como ponto de partida para a compreensão das diferenças entre SGBDs relacionais e analíticos.
- Date (2004) → reforço conceitual sobre dados, informação, independência de dados e fundamentos dos SGBDs, apoiando a distinção entre sistemas transacionais (OLTP) e analíticos (OLAP).
- Laudon & Laudon (2014) → contextualização organizacional da informação, destacando o papel estratégico dos sistemas de informação e dos ambientes de Data Warehouse no apoio à tomada de decisão.
- Inmon (2005) → fundamentação clássica de Data Warehouse corporativo, com foco em integração, orientação por assunto, variabilidade temporal e não volatilidade dos dados.
- Kimball & Ross (2013) → base conceitual e prática para modelagem dimensional (fatos e dimensões), amplamente utilizada em projetos de DW modernos e alinhada ao material anexo trabalhado.
- Ribeiro & Santos (2020) → sustentação teórica da evolução do dado para informação e conhecimento, contribuindo para a compreensão do valor analítico dos ambientes de DW, Data Lake e Lakehouse.
- Torino (2023) → fundamentação contemporânea sobre arquitetura de dados, metadados, governança e integração de repositórios analíticos, conectando DW, Data Lake e Lakehouse em uma visão arquitetural atual.

# Obrigado!

