



# Unidade 1

## Bancos de Dados Relacionais e Linguagem SQL

---

# 1.1 Processo de modelagem de dados

---

# Dados vs. Informações

- Atuando no setor de tecnologia da informação, é essencial entender como os dados são modelados e armazenados em um banco de dados.
- Atuando em outro setor, é provável que você tenha que trabalhar com dados armazenados em um algum local: local ou nuvem e utilize dados para criar relatórios e/ou tomar decisões.

# Dados vs. Informações

- Dados e informações, geralmente, são usadas como se fossem sinônimos porém têm significados diferentes:
  - **Dados:** Material bruto ou não processado.
  - **Informação:** conhecimento, inteligência, dado com um significado ou uma função especial. Geralmente, é um resultado como combinar, comparar, analisar ou executar cálculos.

# O que é um banco de dados?

- É um conjunto centralizado e estruturado de dados armazenados em um sistema de computador para fornecer recursos para recuperar, adicionar, modificar e excluir os dados quando necessário.
- Estes recursos possibilitam a transformação destes dados recuperados em informações úteis como calculo de média.

# Tem relação com o meu dia a dia?

- A maior parte dos sites que você visita são habilitados por um banco de dados.
- Os aplicativos de locomoção, mensagens, fotografias, entre outros armazenam dados em bancos de dados.
- **Na era da informação todos estes dados tem valor financeiro!**

# E a história, como foi?

- Na perspectiva atual da tecnologia da informação utilizando-se computador, sistema de videogame ou o smartphone, e assim percebemos a distância percorrida.
- De cartuchos de videogame a IA adaptativa personalizada para o jogador e de telefones de disco, cartões de visita, câmeras fotográficas e de filmagens para **SMARTPHONES!**

# E a história, como foi?

- Resolvemos o problema do armazenamento (mídias, capacidade, velocidade)!
- E a recuperação de dados, como fica? A diferença entre todos os tipos de mídias foi unificada?
- Mas como organizar então uma nova forma de trabalhar?



# E a história, como foi?

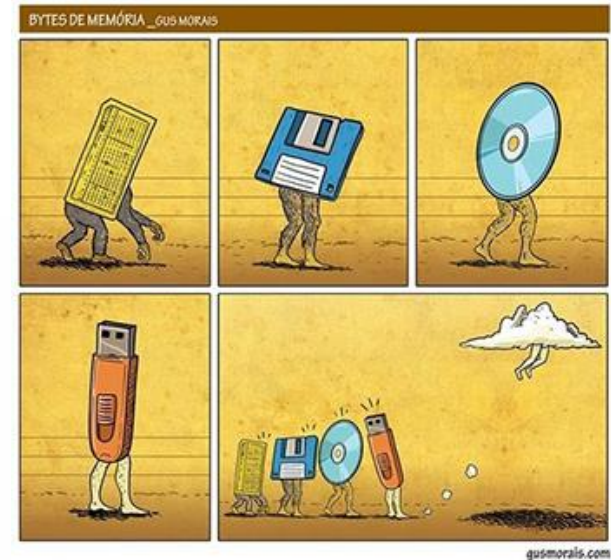


**Fonte:** NETSCAN,2021



**Fonte:** CBL Tech,2021

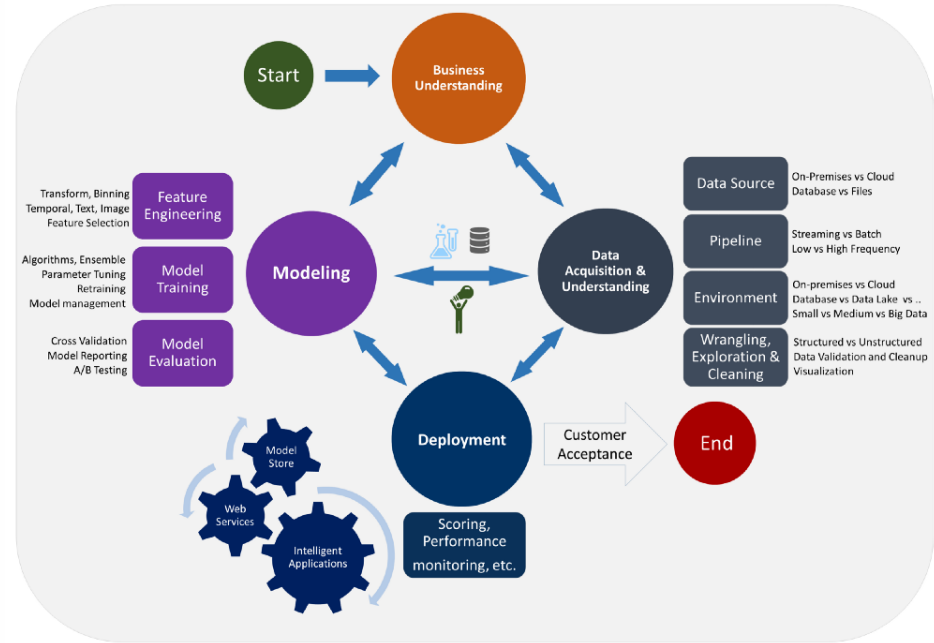
**Fonte:** NETSCAN,2021



# Ciclo de vida da ciência dos dados!



## Data Science Lifecycle

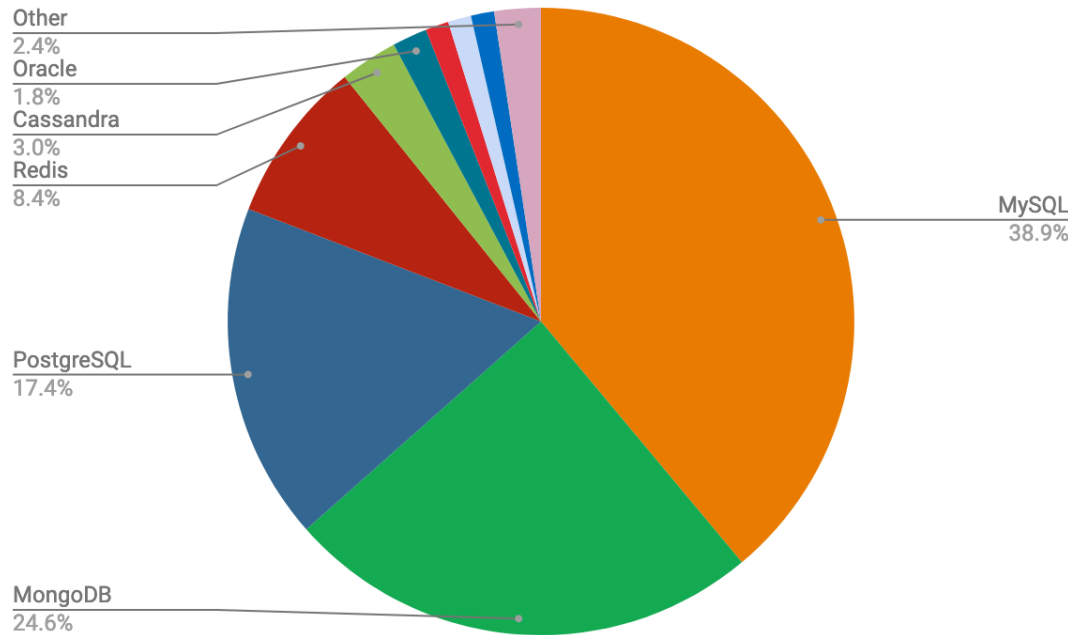


Fonte: TERA, 2020

# Modelagem de dados

- A modelagem de dados é a primeira etapa do desenvolvimento de um banco de dados!
- Envolve as etapas mínimas de coleta, análise e diagramação de dados que resolvem um certo problema de uma empresa ou grupo de pessoas, ou melhor, **atendam a um requisito de negócio (esta é uma linguagem mais formal da área técnica)**.

# Mas porque bancos de dados relacionais?



- MySQL, PostgreSQL, Oracle compõe 60% dos bancos de dados que encontraremos por aí!

# Mas porque bancos de dados relacionais?

- Esta participação significativa nas tendências nos mostra que os bancos relacionais baseados em linguagem SQL terão participação significativa no mercado.
- Além disto, temos o histórico (legado) que se iniciou na década de 60 e ainda sobrevive.

# Processo de modelagem de dados

- Segue algumas etapas básicas:

✓ Projeto conceitual

✓ Projeto lógico

✓ Projeto físico



# Processo de modelagem de dados

- Na prática, podemos considerar os produtos de cada etapa como:
  - ✓ Projeto conceitual: diagrama entidade relacionamento.
  - ✓ Projeto lógico: script de criação das estruturas.
  - ✓ Projeto físico: schemas otimizados nos SGBDs.

# Processo de modelagem de dados

- **Projeto conceitual** está relacionado com o mapeamento dos requisitos, ou seja, mapeamento das variáveis que precisaremos para resolver o problema.
- Representa, através de um diagrama conceitual, a captura precisa das necessidades de dados de um negócio.



# Modelo conceitual

- Busca descrever as necessidades de informação da empresa.
- Cria a documentação do sistema, alimentando de forma sólida o processo de design do banco de dados físico.
- Documenta os processos (regras de negócios) da empresa, considerando as leis e os regulamentos que regem o setor.
- Identifica as entidades e seus relacionamentos.

# Processo de modelagem de dados

- **Projeto lógico** transforma o projeto conceitual em um diagrama equivalente porém **(na prática)**, utiliza-se uma ferramenta case de modelagem.
- Esta automação da modelagem nos permitirá gerar um **script : conjunto de códigos em linguagem específica**, que devidamente importado criará as estruturas em um SGBD.

# Modelo lógico

- Descreve os dados com o maior detalhe possível, sem se preocupar com sua implementação física no banco de dados.
- Deriva de um modelo de dados conceitual porém, inclui todas as entidades e os relacionamentos entre elas ilustrando um DER (Diagrama Entidade Relacionamento).

# Processo de modelagem de dados

- **Projeto físico** além das atividades de criação (**execução dos scripts**) implementa-se toda a política de armazenamento, otimização de acesso aos dados e implementação de estruturas de cache por exemplo.
- Além disto, são criadas as políticas de backup, rotinas de importação e manutenção de índices e etc.

# Modelo Físico

- É uma extensão do modelo lógico onde, define-se precisão, tipos de dados e definições de tabelas.
- Descreve como os objetos devem ser implementados em um banco de dados específico.
- Mostra todas as estruturas de tabelas, incluindo colunas, chaves primárias e chaves estrangeiras

# Finalizando

- A modelagem de dados é a arte de planejar, desenvolver e comunicar que permite a produção de pessoas para atingir um resultado desejado.
- Neste processo captura-se os conceitos e as regras importantes que formatam para mostrá-los visualmente em um diagrama que se torna a base para criar algo físico.

# Processo de modelagem de dados

- Me conte sobre seu conhecimento sobre este assunto:

<https://www.menti.com/a6d8djuq1u>

---

**Let's make it!**

---



# Referências bibliográficas

**SCALE GRID**, 2020. What is Greenplum Database? Intro to the Big Data Database. Disponível em: <https://scalegrid.io/blog/what-is-greenplum-database-intro-to-the-big-data-database/> Acesso em: 18 set. 2021.

**SCALE GRID**, 2019. 2019 Database Trends – SQL vs. NoSQL, Top Databases, Single vs. Multiple Database Use . Disponível em: <<https://scalegrid.io/blog/2019-database-trends-sql-vs-nosql-top-databases-single-vs-multiple-database-use/>> Acesso em: 18 Set. 2021

**SRIVASTAVA**, Shivam. Top 15 Databases to Use in 2021 and Beyond. Disponível em: <<https://appinventiv.com/blog/top-web-app-database-list/>> Acesso em: 18 Set. 2021

# Referências bibliográficas

**CBL Tech, 2021.** Quais são os tipos de armazenamento de dados existentes?. Disponível em: <<https://cbltech.com.br/blog/quais-sao-os-tipos-de-armazenamento-de-dados-existentes.html>>. Acesso em: 19 set. 2021.

**NETSCAN, 2021.** Armazenamento em nuvem: como preservar documentos digitais?. Disponível em: <<https://netscandigital.com/blog/armazenamento-em-nuvem/>>. Acesso em: 21 set. 2021.

**GONÇALVES, Pollyana.** Afinal, como se desenvolve um projeto de Data Science?, 2018. Disponível em: <<https://medium.com/techbloghotmart/afinal-como-se-desenvolve-um-projeto-de-data-science-233472996c34>>. Acesso em: 21 set. 2021.

**TERA, 2020.** Manual prático do ciclo de vida dos dados para iniciantes como eu, 2020. Disponível em: <<https://medium.com/somos-tera/manual-do-ciclo-de-vida-dos-dados-para-iniciantes-67875ff1134d>>. Acesso em: 21 set. 2021