

# Banco de dados

---

PROF.: SALUSTIANO OLIVEIRA

E-MAIL: SALUSTIANO.OLIVEIRA@ESTACIO.BR

# Conteúdo programático

---

1. Introdução e conceitos de banco de dados
  1. Introdução aos sistemas gerenciadores de banco de dados
  2. Componentes e arquitetura de sistemas gerenciadores de banco de dados
  3. Abordagem relacional
2. Modelagem de dados
  1. Conceitos e tipos de modelos de dados
  2. Modelagem conceitual - diagrama entidade relacionamento
  3. Modelagem lógica - modelo relacional de dados
  4. Álgebra relacional

# Conteúdo programático

---

## 3. Normalização

1. Conceitos, princípios e anomalias
2. Formas normais

## 4. SQL

1. Linguagem de definição de dados – DDL
2. Linguagem de manipulação de dados – DML
3. Comando select - expressões e junções
4. Agregando dados
5. Consultas avançadas - (subconsultas e correlatas)
6. Visões

# Conteúdo programático

---

- 5. Indexação & transação
  - 1. Tipos de índices
  - 2. Conceito de transação
  - 3. Propriedades ACID
  - 4. Execução concorrente de transações

# Introdução

---

# Visão Geral

## Dado e Informação

---

- Sociedade do Conhecimento
  - “As atividades que ocupam o lugar central das organizações não são mais aquelas que visam produzir ou distribuir objetos, mas aquelas que produzem e distribuem informação e conhecimento”.

Peter Drucker (1999).

# O que é um dado?

---

- Dados são fatos brutos, ou seja, que não foram submetidos a nenhum processamento de modo a mostrar seu real significado.
- Algo que possa ser representado por um valor.

# O que é uma informação?

---

- Informação é todo o conjunto de dados devidamente ordenados e organizados de forma a terem significado.
- Informação = (dado + significado)
  - Dado → Processado → Informação
  - Informação → Analisada → Conhecimento.
- Os dados e as informações consistem em um dos principais ativos de uma empresa em nossa atualidade.



# Dado, informação e conhecimento

---

**Dados**



**Informação**



**Conhecimento**



# Conceitos

---

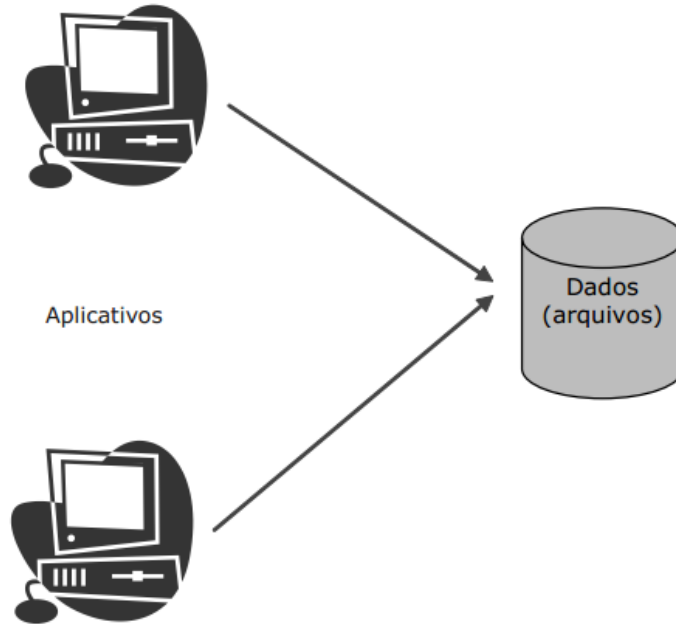
# Sistema de arquivos

---

- Um sistema de arquivos é um conjunto de estruturas lógicas e de rotinas, que permitem ao sistema operacional controlar o acesso ao disco rígido.
- Diferentes sistemas operacionais usam diferentes sistemas de arquivos.
- Existem diversos sistemas de arquivos diferentes, que vão desde sistemas simples como o FAT16, que utilizamos em cartões de memória, até sistemas como o NTFS, EXT4 e APFS, que incorporam recursos muito mais avançados.

# Sistema de arquivos

---



# Problemas do sistema de arquivos

---

- A manutenção é prejudicada pois a estrutura de arquivos é definida e padronizada no próprio código do aplicativo.
- O compartilhamento de um arquivo por vários programas apresenta dificuldades para gerenciar o acesso a esses arquivos e seu controle.
- O desenvolvimento de arquivos e programas de um mesmo SO é realizado isoladamente por programadores e linguagens diferentes.
- A falta de gerenciamento de acessos concorrentes aos dados e recuperação de dados.

# Banco de dados

---

- Banco de dados é uma coleção de dados inter-relacionados.
- Um banco de dados nada mais é que uma coleção de dados persistentes, dados estes que são utilizados por aplicações empresariais.
- Construído em atendimento a uma proposta específica.

# Sistema gerenciador de banco de dados

---

- Sistema Gerenciador de Banco de Dados (SGBD) é um conjunto de programas e ferramentas utilizadas para configurar, atualizar e manter um banco de dados.
  - Recursos para administrar usuários/permisões.
  - Recursos para criar/alterar tabelas e banco de dados.
  - Recursos para backup e restauração de dados.
  - Recursos para otimizar a performance do banco.

# Banco de dados

---





# Modelo de dados

---

- Modelo de dados é coleção de ferramentas para descrever os dados, seus relacionamentos, semântica e restrições de consistência.
  - Modelo relacional
  - Entidade-relacionamento (ER)
  - Baseado em objetos
  - Semiestruturado (XML)

# Modelo relacional

---

- Coleção de tabelas para representar dados e os relacionamentos entre eles.
- Cada tabela contém registros de um mesmo tipo
- Cada registro define um número fixo de campos ou atributos
- Modelo mais usado atualmente

# Entidade-relacionamento (ER)

---

- Baseado na percepção do mundo real que consiste em uma coleção de objetos (entidades) e os relacionamentos entre esses objetos.

# Baseado em objetos

---

- Pode ser visto como uma extensão do modelo ER com noções de encapsulamento e identidade do objeto.

# Semiestruturado (XML)

---

- Permite que registros do mesmo tipo possuam diferentes atributos – oposto daquilo que existe nos outros modelos.

# Esquema

---

- Descrição de uma coleção de dados em particular com base em um modelo de dados.
- O termo “esquema de banco de dados” refere-se a uma representação visual de um banco de dados, um conjunto de regras que governa um banco de dados ou todo o conjunto de objetos pertencentes a um determinado usuário.

# Banco de dados relacional

---

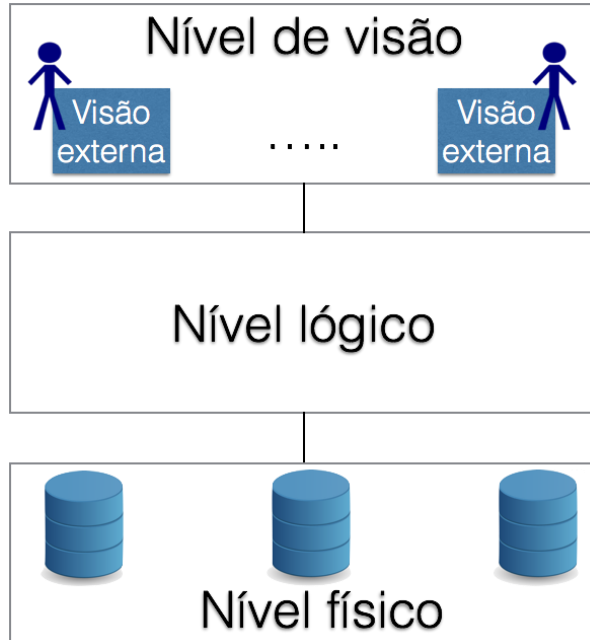
Tabela Alunos

Id	Nome	Curso
123	Juca	Telecomunicações
456	Amélia	Elétrica
789	Breno	Computação
900	Jucé	Automação

Tabela Cursos

Curso	Campus	Orçamento
Telecomunicações	São José	100.000,00
Elétrica	Florianópolis	300.000,00
Computação	Lages	20.000,00
Automação	Chapecó	150.000,00

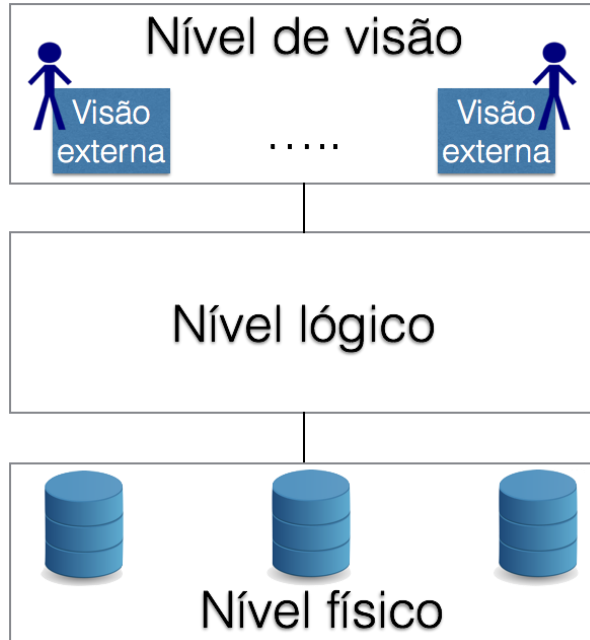
# Abstração e independência de dados



- **Nível físico** – Descreve como os dados são armazenados em baixo nível.
- Tabelas consistem em um conjunto de arquivos, cujo conteúdo não possui qualquer ordenação
- Índices são dados específicos que possuem uma ordenação

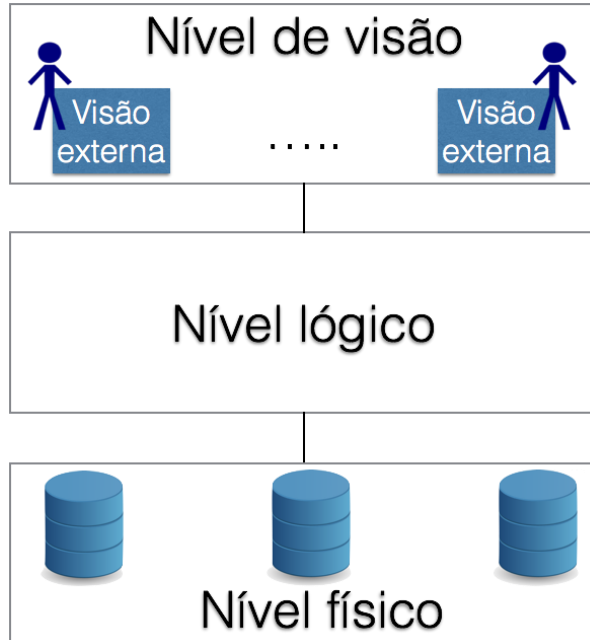


# Abstração e independência de dados



- **Nível lógico** – Descreve quais dados são armazenados e os relacionamentos entre eles.
  - Ex.: Tabela Aluno possui id, nome do aluno e nome do curso.

# Abstração e independência de dados



- **Nível de visão** – esquema externo que descreve a parte do banco de dados que um grupo de usuários pode ver
  - Ex.: O coordenador de curso só consegue ver informações sobre os alunos de seu curso

# Linguagem de definição de dados

## *Data Definition Language - (DDL)*

---

- Linguagem para especificação do esquema ou estrutura do banco de dados.
- Compilador DDL gera um conjunto de tabelas em um conjunto de arquivos denominado **dicionário de dados**.

```
CREATE TABLE Alunos(  
  Id      INT AUTO_INCREMENT,  
  Nome    VARCHAR(60) not null,  
  Curso   VARCHAR(60) null,  
  PRIMARY KEY(Id)  
);
```

# Linguagem de manipulação de dados

## *Data Manipulation Language - (DML)*

---

- Permite acessar ou manipular dados no banco de dados.
  - Recuperação, Inserção, Exclusão ou Modificação
- **SQL** (*Structured Query Language*) é a linguagem de consulta mais utilizada atualmente.

```
SELECT Nome
FROM Alunos
WHERE Alunos.Id = '123';
```

```
SELECT Alunos.Nome, Alunos.Curso, Cursos.Campus
FROM Alunos, Cursos
WHERE Alunos.Curso = Cursos.Curso
AND Cursos.Curso = "Telecomunicacoes";
```

# Existe algum problema com esse projeto?

Id	Nome	Curso	Disciplina	Professor
123	Juca	Telecomunicações	Sinais	João
456	Amélia	Elétrica	Sinais	João
789	Breno	Computação	Programação I	Martin
900	Jucé	Automação	Programação I	Martin
334	Maira	Telecomunicações	Sistemas Distribuídos	Paulo
453	Célio	Telecomunicações	Cálculo II	Luíza
112	Cícero	Computação	Cálculo II	Luíza
322	Marco	Automação	Cálculo II	Luíza
567	Alonso	Computação	Sistemas Distribuídos	Paulo
257	Luiz	Telecomunicações	Sinais	João

# Normalização

---

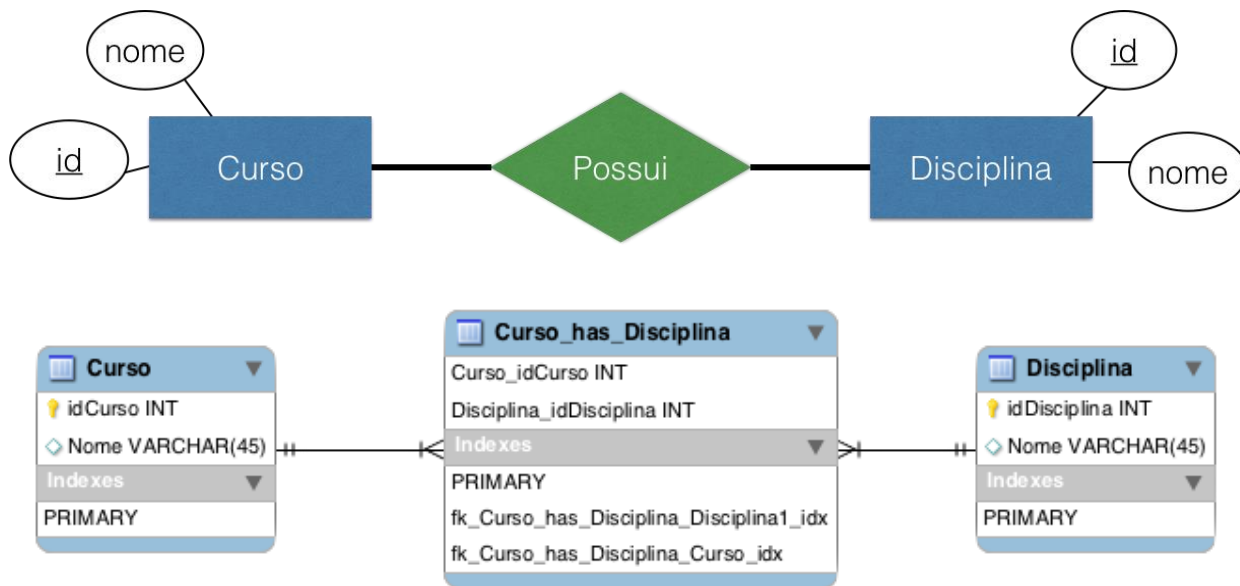
- Normalização é o processo de modelar o banco de dados projetando a forma como as informações serão armazenadas a fim de eliminar, ou pelo menos minimizar, a redundância no banco.
- Tal procedimento é feito a partir da identificação de uma anomalia em uma relação, decompondo-as em relações melhor estruturadas.

# Diagramas ER

---

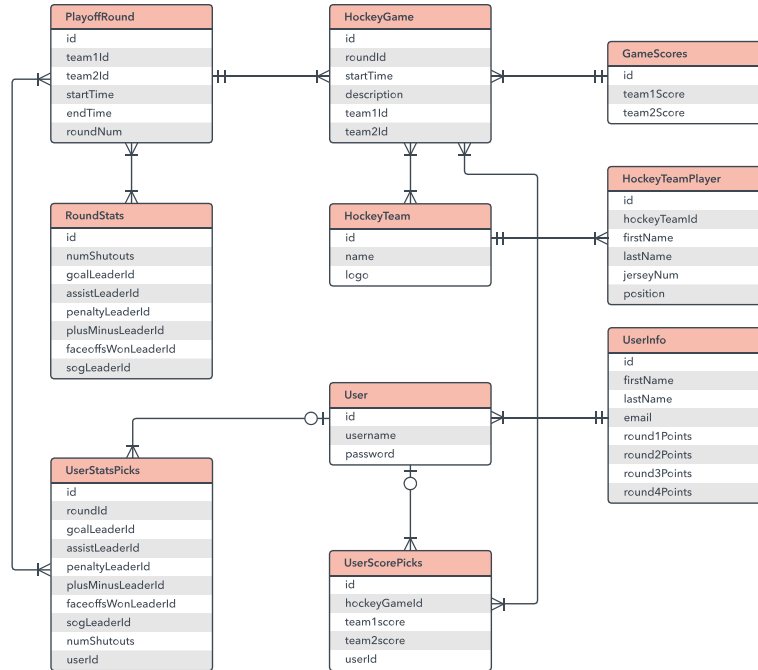
- Um diagrama entidade relacionamento (ER) é um tipo de fluxograma que ilustra como “entidades”, p. ex., pessoas, objetos ou conceitos, se relacionam entre si dentro de um sistema.
- Diagramas ER são mais utilizados para projetar ou depurar bancos de dados relacionais nas áreas de engenharia de software, sistemas de informações empresariais, educação e pesquisa.

# Diagramas ER





# Diagramas ER



# Banco de dados transacional

## *Propriedades ACID*

---

- Garante que todas operações de consulta ou de alteração são atômicas, consistentes, isoladas e duráveis.
  - Atomacidade
    - Todas operações (leitura/escrita) em uma transação são executadas com sucesso ou tudo é desfeito.
  - Consistência
    - A execução de uma transação leva o banco de um estado consistente para um outro estado consistente.
  - Isolamento
    - Transações podem acontecer de forma concorrente sem qualquer interferência.
  - Durabilidade
    - Ao concluir uma transação, todas modificações geradas serão persistentes.

# Transação

---

```
read saldoContaOrigem  
write (saldoContaOrigem - 500)  
read saldoContaDestino  
write (saldoContaDestino + 500)
```

- **Antes**

- Saldo origem: 600,00
- Saldo destino: 300,00

- **Depois**

- Saldo origem: 100,00
- Saldo destino: 800,00

# Transação

---

```
read saldoContaOrigem  
write (saldoContaOrigem - 500)  
read saldoContaDestino  
write (saldoContaDestino + 500)  
FALHA NO SISTEMA
```

- **Antes**

- Saldo origem: 600,00
- Saldo destino: 300,00

- **Depois**

- Saldo origem: 100,00
- Saldo destino: 800,00

# Transação

---

```
read saldoContaOrigem  
write (saldoContaOrigem - 500)
```

FALHA NO SISTEMA

```
read saldoContaDestino  
write (saldoContaDestino + 500)
```

- **Antes**

- Saldo origem: 600,00
- Saldo destino: 300,00

- **Depois**

- Saldo origem: 100,00
- Saldo destino: 300,00

# Transação

---

- SGBD's garantem que a execução concorrente das transações  $T_1, \dots, T_n$  seja equivalente a uma execução serial dessas transações.
  - Antes de ler ou escrever, a transação obtém acesso a uma seção crítica e só libera depois que for concluída.
- A atomicidade e durabilidade podem ser garantidas por meio de um arquivo de registro (log).
  - Mantenha no log todas as escritas que foram concluídas.
  - Se houver alguma falha, então desfaça as escritas parciais (que não foram escritas no log).

# Alguns SGBD's relacionais

---



PostgreSQL

ORACLE®  
DATABASE



MySQL®



Microsoft®  
SQL Server®



# Big data

---

- Grande volume de dados.
- Gerados em grande velocidade.
- Com grande variedade.
- Exigem novas formas de processamento para ajudar nas tomadas de decisão.
- Fundamentado sobre tecnologias de análise (Hadoop / MapReduce) e infraestrutura de armazenamento e processamento.

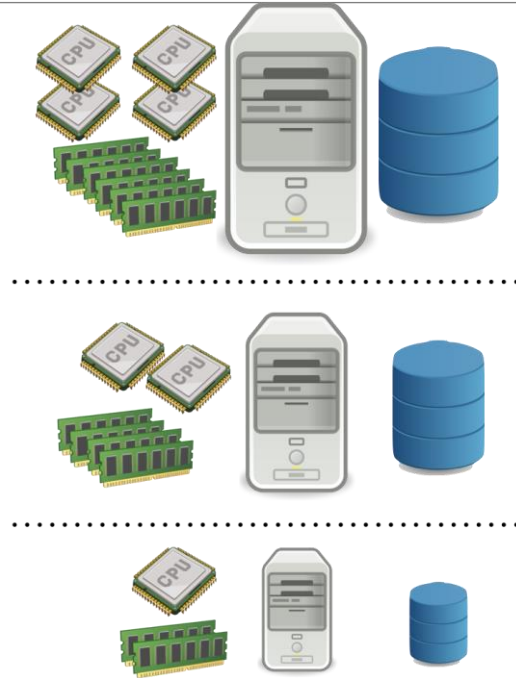


# Banco de dados NoSQL

---

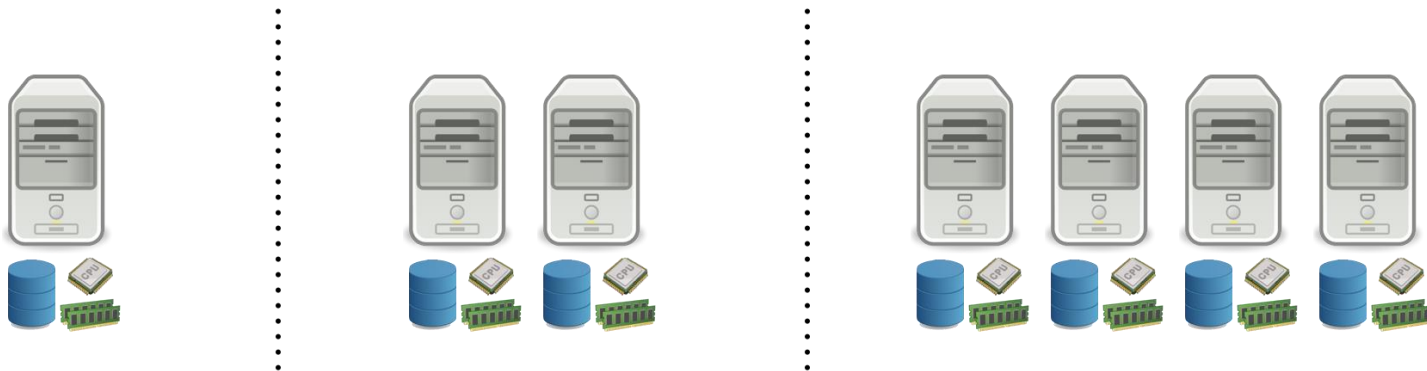
- Adequado para armazenar uma vasta quantidade de dados de maneira efetiva, com baixo custo e com facilidade para implementar escalonamento horizontal (distribuição por diversos nós).
  - Modelo de dados não é baseado em tabelas
    - Em uma tabela, todas as linhas terão o mesmo número de colunas
  - Estrutura de dados em bancos **NoSQL**
    - chave-valor
    - colunas esparsas
    - grafo
    - orientado a documento

# Escalonamento vertical



- Adiciona mais recursos (memória, cpu, disco) em um único nó para permitir que atenda o aumento da demanda.

# Escalonamento horizontal



- Capacidade de processamento e armazenamento é aumentada por meio de adição de novos nós no cluster.
- Escalabilidade horizontal é a opção mais barata e flexível, porém possui uma maior complexidade.

# Encerramento

---

- Definimos um banco de dados como uma coleção de dados relacionados, na qual os dados significam fatos registrados.
- Um típico banco de dados representa alguns aspectos da vida do mundo real e é utilizado por um ou vários grupos de usuários para propostas específicas.
- Um SGBD é um pacote de software para a implementação e manutenção de bancos de dados computadorizados.

# Exercícios

---

- Defina os seguintes termos: dados, banco de dados e SGBD.
- Quais as características principais da abordagem de um banco de dados e como ela se difere dos sistemas tradicionais de arquivos.
- Quais são os diferentes tipos de usuários finais de banco de dados?
- Quais as funcionalidades que podem ser fornecidas por um SGBD.