

TI TOTAL

ÁREA FISCAL E CONTROLE



 Professor
Ramon Souza

Tecnologia da Informação

RESUMO

Bancos de Dados

CONCEITOS PRELIMINARES

- **Um banco de dados (BD):**
 - Representa algum **aspecto do mundo real** (minimundo ou universo de discurso).
 - É uma **coleção logicamente coerente de dados** com significado inerente.
 - É projetado, construído e populado com dados para **finalidade específica**.
- Um **sistema gerenciador de banco de dados (SGBD)** é uma **coleção de programas que permite aos usuários criar e manter um banco de dados**.
 - **Definição:** especificar os tipos, estruturas e restrições dos dados.
 - **Construção:** armazenar os dados em algum meio controlado.
 - **Manipulação:** consulta, atualização e geração de relatórios.
 - **Compartilhamento:** acesso simultâneo por usuários e sistemas.
 - **Proteção:** contra defeitos e acesso não autorizado ou malicioso.
 - **Manutenção:** evolução do sistema ao longo do tempo.
- **Sistema de banco de dados (SBD) = BD + SGBD.**

CARACTERÍSTICAS DOS BANCOS DE DADOS

As principais características de um banco de dados são:

- **Natureza de autodescrição:** contém uma **definição ou descrição completa de sua estrutura e restrições**, além dos dados em si.
- **Abstração de dados:** SGBD oferece aos usuários **uma representação conceitual de dados que não inclui muitos detalhes** de como os dados são armazenados ou como as operações são implementadas.
 - **Independência dos dados dos programas:** **estrutura dos arquivos de dados** é armazenada no catálogo do SGBD **separadamente dos programas** de acesso.
 - **Independência da operação do programa:** os programas podem atuar invocando as operações, **independentemente de como estas estão implementadas**.
- **Suporte de múltiplas visões de dados:** cada usuário pode exigir um **ponto de vista ou visão diferente do banco de dados**.
- **Compartilhamento de dados e processamento de transação multiusuário:** o SGBD precisa incluir um software de **controle de concorrência** para **garantir que vários usuários tentando atualizar o mesmo dado faça isso de maneira controlada**.

TRANSAÇÕES ACID

As propriedades das transações são:

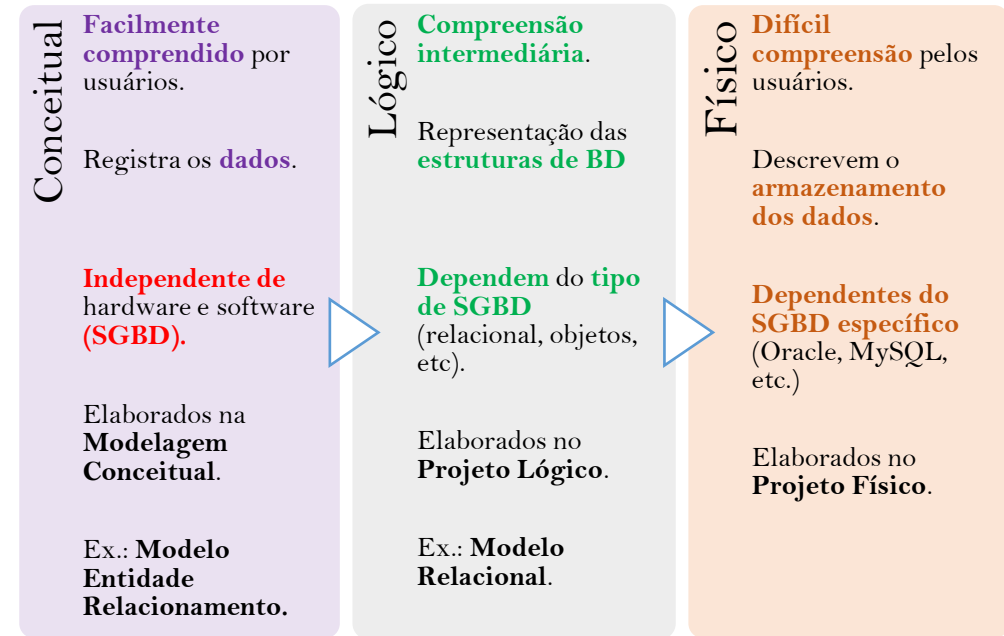
| | |
|---------------------|--|
| Atomicidade | Conceito: unidade de processamento atômica. Transação deve ser realizada em sua totalidade ou não deve ser realizada de forma alguma . |
| Consistência | Conceito: transação deve levar o banco de um estado consistente para outro . Regras e restrições respeitadas . |
| Isolamento | Conceito: transação deve parecer executar isolada das demais . Não interferência entre transações paralelas. |
| Durabilidade | Conceito: mudanças realizadas devem ser persistidas no banco de dados . Mudanças não devem ser perdidas por falhas posteriores. |

METADADOS E CATÁLOGO DE DADOS

Metadados (informações do descritor) são dados estruturados que **descrevem, identificam, explicam, localizam** e, portanto, facilitam a recuperação, uso e gestão de recursos de informação. São os chamados **dados sobre outros dados**.

Catálogo ou dicionário de dados é o **local onde ficam armazenados os metadados**. Contém informações como **a estrutura de cada arquivo, o tipo e o formato de armazenamento** de cada item de dados e diversas **restrições** sobre os dados.

MODELOS DE BANCOS DE DADOS



ESQUEMA x INSTÂNCIA

Um **esquema** de banco de dados representa a **estrutura do banco**, mas **sem os dados**. Em outra definição, é o **conjunto de regras que governa um banco** de dados.

Uma **instância** de um banco de dados representa os **dados** que estão **armazenados em um dado instante**.

ARQUITETURA DE TRÊS ESQUEMAS

A arquitetura de três esquemas visa separar as aplicações do usuário do banco de dados físico e possui **três níveis**:

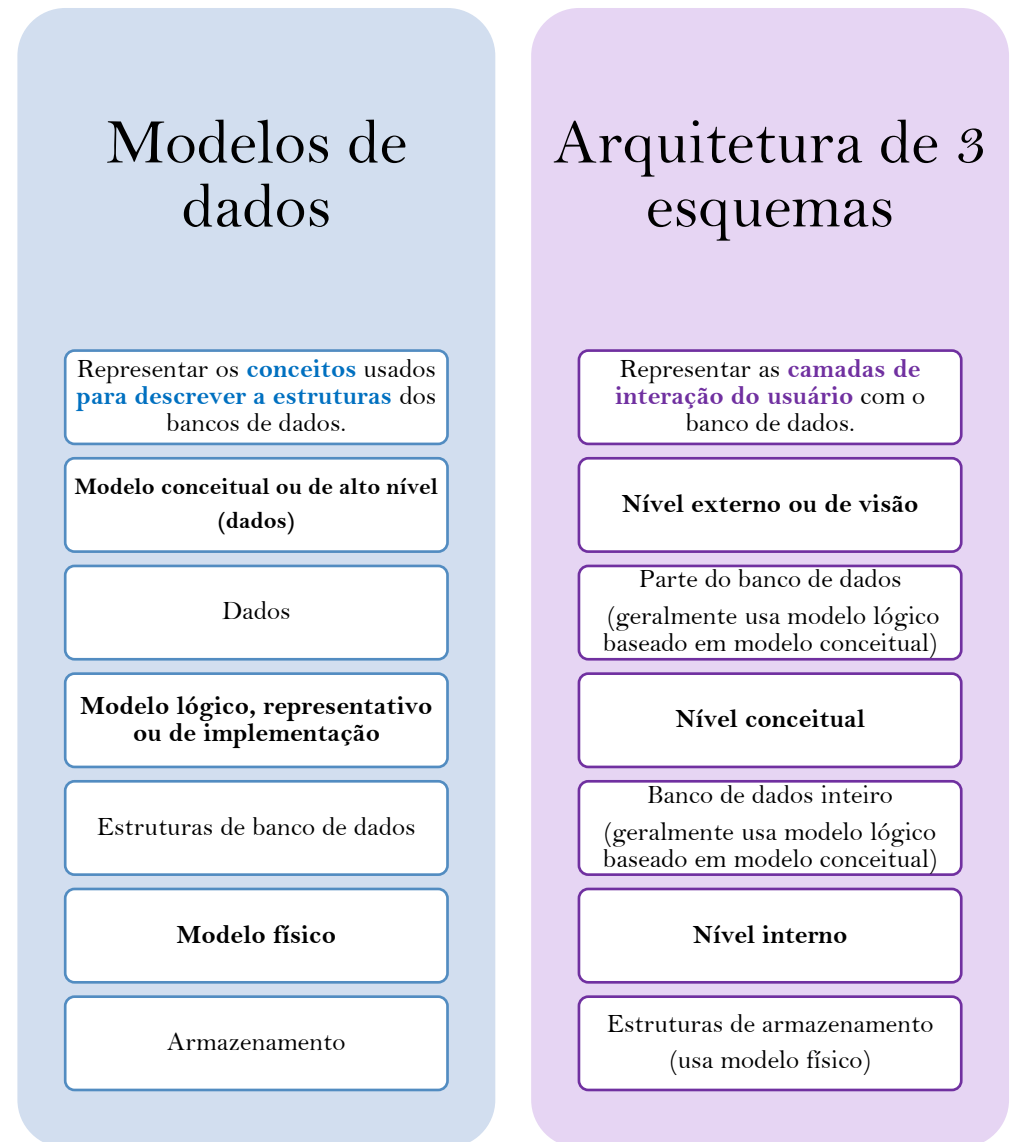
- **Nível externo ou de visão:** **descreve a parte do banco de dados em que um usuário em particular está interessado** e **oculta o restante do banco de dados** do grupo de usuários.
- **Nível conceitual:** **descreve a estrutura do banco de dados inteiro para uma comunidade de usuários**. O esquema conceitual **oculta os detalhes das estruturas de armazenamento físico**.
- **Nível interno:** **descreve a estrutura do armazenamento físico do banco de dados**.

INDEPENDÊNCIA LÓGICA E FÍSICA

A independência pode ser:

- **Independência lógica de dados:** **capacidade de alterar o esquema conceitual sem ter de alterar os esquemas externos ou de programas de aplicação**.
- **Independência física de dados:** **capacidade de alterar o esquema interno sem ter de alterar o esquema conceitual** e, por consequência, sem ter que alterar os esquemas externos.

MODELOS x ARQUITETURA



PRINCIPAIS MODELOS DE DADOS

Os principais modelos de dados são:

- **Modelo relacional:** banco de dados como uma **coleção de tabelas** bidimensionais (linhas x colunas).
- **Modelo de objetos:** banco de dados em termos de **objetos, suas propriedades (atributos) e operações (métodos)**. Termos relacionados: herança (subclasse e superclasse), encapsulamento e polimorfismo.
- **Modelo objeto-relacional:** **SGBDs relacionais** estendidos para **incorporar conceitos de bancos de dados de objeto**.
- **Modelo XML:** utiliza **estruturas de árvore hierárquicas** com uso de tags (etiquetas).
- **Modelo de rede:** modelo legado (mais antigo) que representa os **dados como tipos de registro e também representa um tipo limitado de relacionamento 1:N**, chamado de tipo de conjunto.
- **Modelo hierárquico:** representa os dados como **estruturas de árvore hierárquicas**. Cada hierarquia simboliza uma série de registros relacionados em uma estrutura de **“pais” e “filhos”**.

MODELO RELACIONAL

| | | |
|--------------------------|---------------------------------|--|
| Modelo Relacional | Conceitos | Conjuntos de relações ou tabelas que representam dados e relacionamentos . |
| | Princípio da Informação | Base teórica reside na teoria de conjunto e lógica de predicado de primeira ordem. Todo o conteúdo de informação do banco de dados é representado como valores explícitos em posições de colunas em linhas de tabelas . |
| | Componentes (ANSI/SPARC) | Uma coleção de estruturas de dados (relações ou tabelas) compondo o nível conceitual. Uma coleção dos operadores , a álgebra e o cálculo relacionais, que constituem a base da linguagem SQL. Uma coleção de restrições da integridade , definindo o conjunto consistente de estados de base de dados e de alterações de estados. |

DOZE (TREZE) REGRAS DE CODD

Doze (treze) regras de Codd para Bancos Relacionais

- Regra 0 - Regra fundamental ou regra base (capacidade exclusivamente relacional)
- Regra 1 - A regra da informação ou Princípio da Informação
- Regra 2 - A regra de acesso garantido
- Regra 3 - Tratamento sistemático de valores nulos
- Regra 4 - Catálogo on-line baseado no modelo relacional
- Regra 5 - Sublinguagem Ampla de Dados
- Regra 6 - Atualização de visões
- Regra 7 - Inserção, Atualização, e Exclusão de alto nível
- Regra 8 - Independência Física de dados
- Regra 9 - Independência Lógica de Dados
- Regra 10 - Independência de Integridade
- Regra 11 - Independência de Distribuição
- Regra 12 - A não-transposição das regras

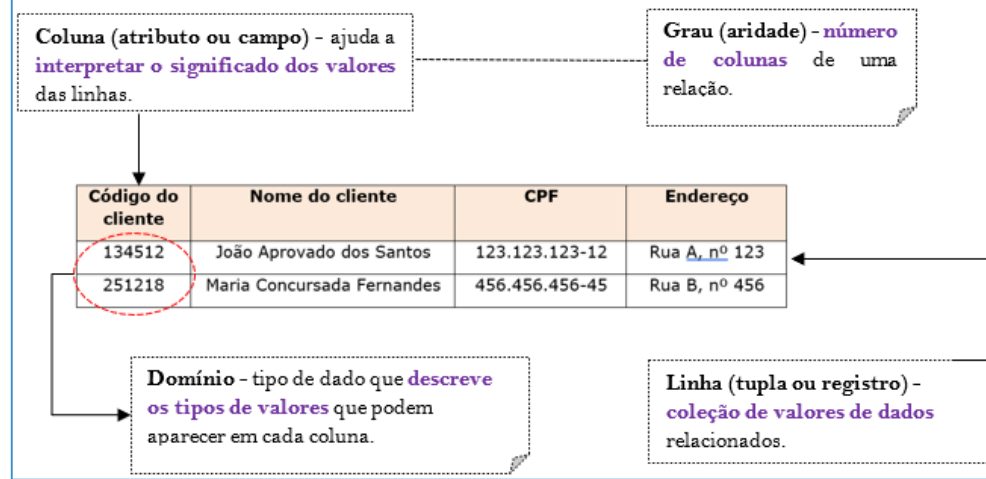
ÁLGEBRA RELACIONAL

Operações da álgebra relacional

| | |
|---------------------------------|---|
| Seleção (σ) | Retorna as tuplas da relação que satisfazem um predicado . (PRIMITIVA) |
| Projeção (Π) | Retorna os atributos especificados de todas as linhas da relação de entrada , removendo as tuplas duplicadas da saída. (PRIMITIVA) |
| Junção (\bowtie) | Retorna pares de linhas das duas relações de entrada que têm o mesmo valor em todos os atributos que possuem o mesmo nome. |
| Produto cartesiano (\times) | Retorna todos os pares de linhas das duas relações de entrada (independentemente de ter ou não os mesmos valores em atributos comuns). (PRIMITIVA) |
| União (\cup) | Retorna as tuplas que estão na primeira relação, na segunda relação, ou em ambas . (PRIMITIVA) |
| Interseção (\cap) | Retorna as tuplas que aparecem tanto na primeira relação quanto na segunda . |
| Diferença ($-$) | Retorna tuplas que aparecem na primeira relação, mas não estão na segunda . (PRIMITIVA) |

CONCEITOS DO MODELO RELACIONAL

Tabela (relação) – representa **dados e relacionamentos**.



VISÃO (VIEW)

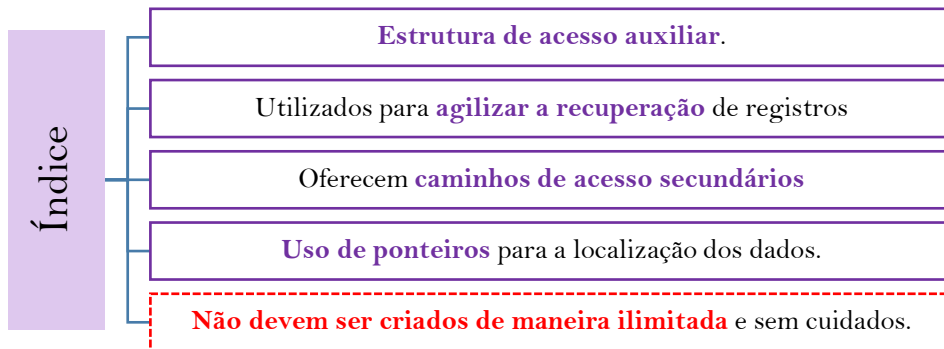
Visão (view)

- Tabela derivada de outras tabelas (tabelas base ou outras visões).
- Maneira alternativa de visualização dos dados de uma ou mais tabelas.
- Consulta pré-definida ou armazenada.
- Tabela virtual: não necessariamente existe em forma física.
- Em regra, não atualizáveis.

Visão materializada (materialized view).

- Visão armazenada.
- Em regra, atualizáveis.

ÍNDICE (INDEX)

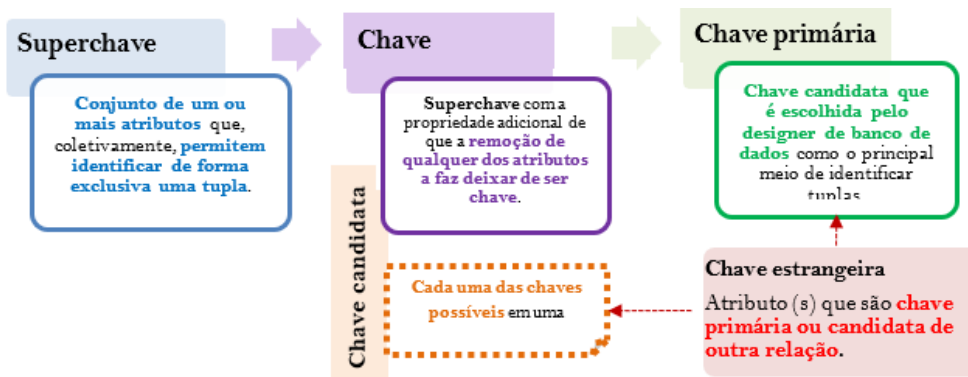


RELACIONAMENTOS

Os relacionamentos podem ser:

- **Relacionamento um para um (1:1):** neste tipo de relacionamento, **cada elemento de uma tabela se relaciona a um único elemento de outra tabela**. Em banco relacional pode ser implementado com:
 - Chave estrangeira em uma das tabelas (modo mais comum).
 - Tabela única mesclada.
 - Terceira tabela para cruzamento das chaves.
- **Relacionamento um para muitos (1:N):** **cada elemento de uma tabela r1 pode ser relacionar com mais de um elemento da tabela r2, no entanto cada elemento da tabela r2 está relacionado a apenas um elemento de r1**. Em banco relacional é implementado:
 - Chave estrangeira no lado muitos da relação.
- **Relacionamento muitos para muitos (N:M):** neste tipo de relacionamento, **para cada valor do campo de uma tabela, pode haver N valores no campo da outra tabela e vice-versa**. Em banco relacional é implementado:
 - Terceira tabela para cruzamento das chaves (tabela de ligação).

CHAVES



A propriedade da identificação de uma tupla pode ser atribuída a qualquer uma das seguintes chaves: superchave, chave, chave candidata, chave primária, chave secundária. Uma (qualquer uma delas) é um atributo ou conjunto de atributos que permite identificar uma tupla. A única chave que não serve para identificação é a estrangeira, pois serve para associação.

NORMALIZAÇÃO

A normalização de dados visa a (1) **minimização de redundância** e (2) a **minimização de anomalias de inserção, exclusão e atualização**.

As definições das formas normais são:

1FN

- Todos os atributos são **atômicos**.
- Não possui atributos multivalorados nem compostos.

2FN

- Está na 1FN.
- Cada atributo não-chave é dependente da chave primária (ou candidata) inteira. (**Dependência funcional total**)
- Cada atributo não-chave não é dependente de apenas parte da chave.

3FN

- Está na 2FN.
- Cada atributo não chave **não possui dependência transitiva**, para cada chave candidata.
- Todos os atributos dependem exclusivamente da chave primária da tabela.

FNBC

- Está na 3FN (variação forte da 3FN).
- Todo **determinante é chave candidata**.
- Não há **dependências entre atributos não chave**.

4FN

- Está na 3FN.
- **Não possui dependência multivalorada**.

5FN

- Está na 4FN.
- **Não possui dependência de junção**.

TEMA AVANÇADO: AXIOMAS

| Axioma | Descrição |
|---------------------------------------|---|
| Reflexibilidade | Se X contém Y, então $X \rightarrow Y$. |
| Expansibilidade ou aumentativa | Se $X \rightarrow Y$, então $XZ \rightarrow YZ$ para qualquer Z. |
| Transitividade | Se $X \rightarrow Y$, e $Y \rightarrow Z$ então $X \rightarrow Z$. |
| União ao aditiva | Se $X \rightarrow Y$, $X \rightarrow Z$ então $X \rightarrow YZ$. |
| Decomposição ou separação | Se $X \rightarrow YZ$, então $X \rightarrow Y$ e $X \rightarrow Z$. |
| Pseudo-Transitividade | $X \rightarrow Y$ e $YZ \rightarrow W$, então $XZ \rightarrow W$ |
| Acumulação | Se $X \rightarrow Y$, então $XZ \rightarrow Y$ |
| Unicidade | Se a: $X \rightarrow Y$ e b: $X \rightarrow Y$; $a=b$ |