

Banco de Dados I

Professor Msc. Aparecido Vilela Junior
aparecido.vilela@unicesumar.edu.br

- Elmasri & Navathe – Fundamentos de Bancos de Dados
- Carlos Alberto Heuser – Projeto de Banco de Dados
- Korth e Silberchatz – Sistema de Bancos de Dados

- Um banco de dados é uma coleção de dados relacionados.
- Os dados são fatos que podem ser gravados e que possuem um significado implícito.
 - Por exemplo, considere nomes, números telefônicos e endereços de pessoas que você conhece.
- Possui as seguintes propriedades implícitas:
 - Um banco de dados representa alguns aspectos do mundo real, sendo chamado, às vezes, de **minimundo ou de universo de discurso (UoD)**. As mudanças no minimundo são refletidas em um banco de dados.
 - Um banco de dados é uma coleção lógica e coerente de dados com algum significado inerente. Uma organização de dados ao acaso (randômica) não pode ser corretamente interpretada como um banco de dados.
 - Um banco de dados é projetado, construído e povoado por dados, atendendo a uma proposta específica. Possui um grupo de usuários definido e algumas aplicações preconcebidas, de acordo com o interesse desse grupo de usuários.

- Os Bancos de Dados fazem parte do nosso dia-a-dia:
 - operação bancária
 - reserva de hotel
 - matrícula em uma disciplina da universidade
 - cadastro na vídeo locadora

- **Dado**: fato do mundo real que está registrado
 - exemplos: endereço, data
- **Informação**: fato útil que pode ser extraído direta ou indiretamente a partir dos dados
 - exemplos: endereço de entrega, idade
- **Banco de Dados (BD)**: coleção de dados inter-relacionados e persistentes que representa um sub-conjunto dos fatos presentes em um domínio de aplicação(universo de discurso)

Por que usar BD?

Considere o contexto de uma grande organização que NÃO utiliza BD

- exemplo: domínio da Universidade
 - várias divisões gerenciais (com suas aplicações)
 - grande volume de dados
 - aplicações manipulam dados comuns

Acadêmica

Alunos
Professores
Disciplinas
Turmas
Salas

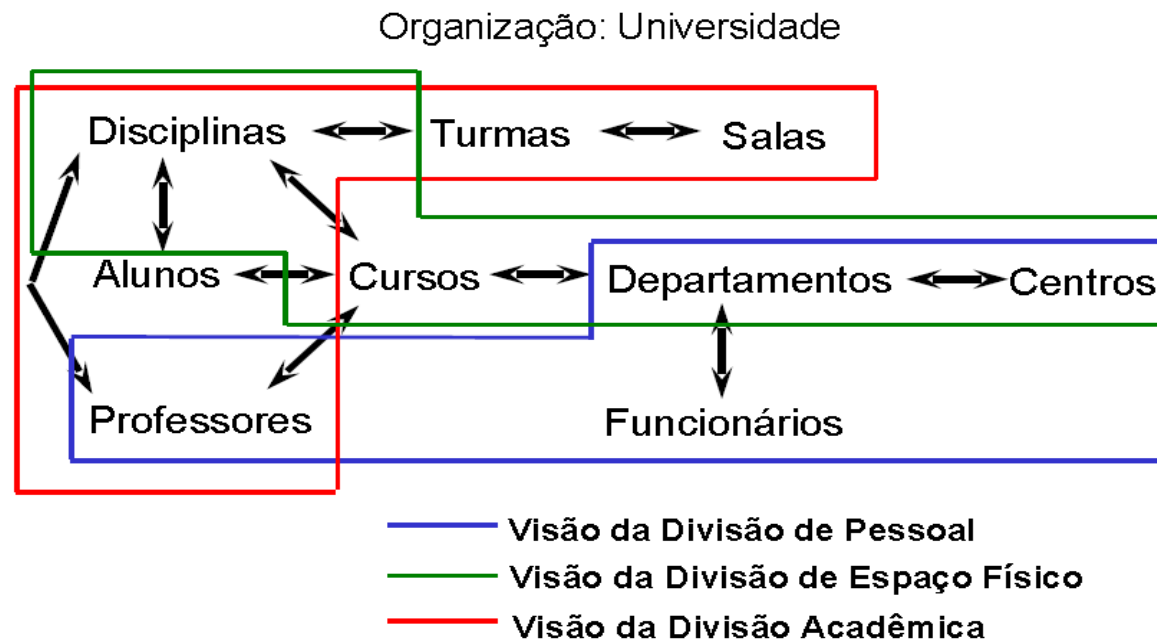
Espaço Físico

Centros
Departamentos
Cursos
Disciplinas

Pessoal

Centros
Departamentos
Professores
Funcionários

Exemplo de um BD



Banco de dados

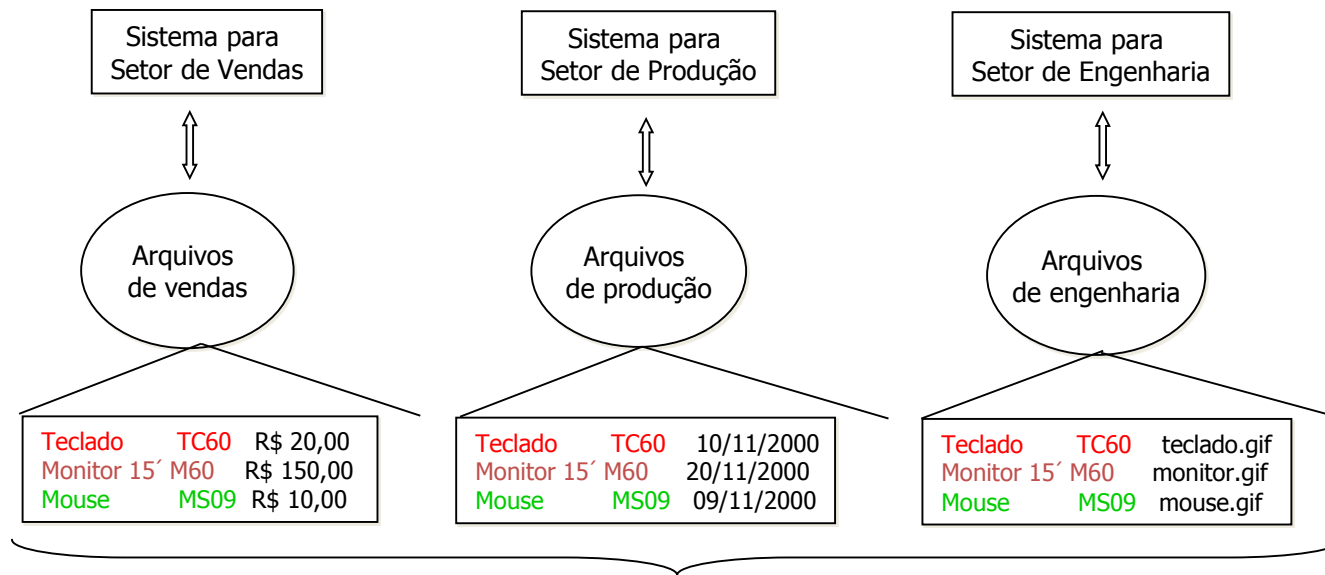
- Banco de dados = instância de dado + meta-dados
 - ✓ Instância de dado
 - Dado propriamente
 - ✓ Meta-dados
 - *Dicionário de dados*
 - Esquema da base de dados
 - Acessado através de linguagens de definição de dados

Nem sempre foi assim...

- Sistemas de Arquivos (armazenados em pastas, no disco):
 - ✓ Funcionalidades oferecidas
 - Registros de tamanho fixo com campos de tipos diferentes
 - Possibilidade de memória virtual e persistência
 - Índices: *hash*, árvore-B
 - Bloqueio de arquivo e registro para concorrência
- Dados de diferentes aplicações **não estão integrados**
- Dados são projetados para atender **uma aplicação específica**

Sistemas de arquivos

- Em uma fábrica com os dados em sistemas de arquivos:



Mesmos dados aparecem em todos os arquivos da fábrica
[baseado em Heuser]

Sistemas de arquivos

dados não integrados

- Mesmo objeto da realidade é representado várias vezes na base de dados
 - ✓ Exemplo - teclado, monitor e mouse

- **Redundância não controlada** de dados
 - ✓ Não há gerência automática da redundância
 - ✓ Redundância leva a
 - *inconsistência dos dados*
 - *re-digitação de informações*
 - *dificuldade de extração de informações*

- Dados **pouco confiáveis** e de **baixa disponibilidade**

Sistemas de arquivos

- Concorrência
 - ✓ Difícil implementação
 - ✓ Políticas de acesso concorrente consistente são independentes de domínio
- Tolerância a falhas
 - ✓ Falta de luz, erro de disco, interrupção de funcionamento, etc
 - ✓ Cópias? restauração do estado anterior? Consistência da base?
- Segurança
 - ✓ Acesso diferenciado por tipo de usuário

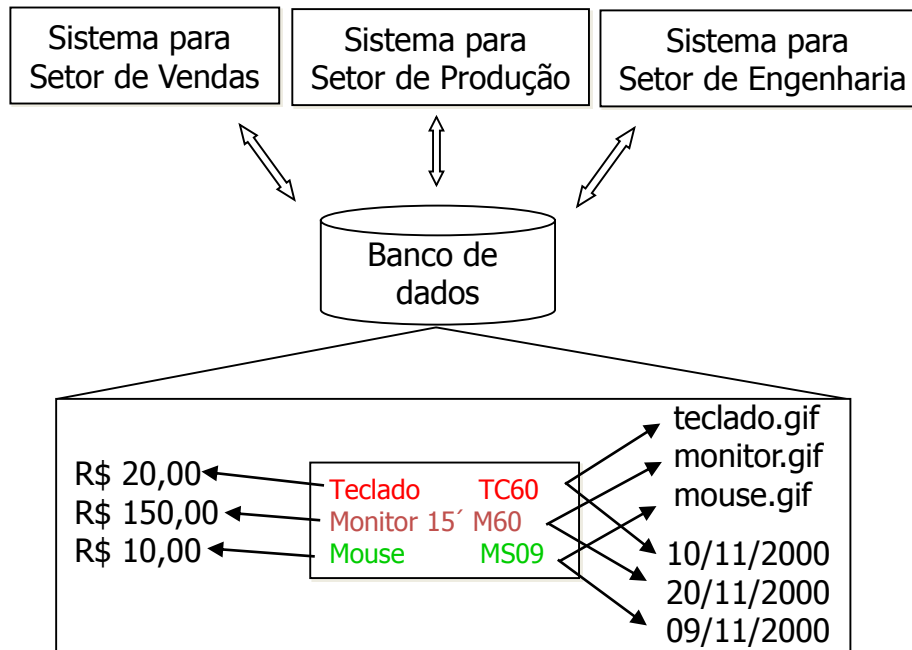
Sistemas de arquivos

gerenciamento dos arquivos

- Outros problemas:
 - ✓ Número máximo de arquivos
 - ✓ Tamanho de memória
 - ✓ Limitações do tipo de arquivo, tipo de acesso
 - ✓ Preocupações técnicas junto com problemas do domínio
- Exemplo: efetuar aluguel de um DVD
 - ✓ Sem reservas? sem multas?
 - ✓ Como registrar um empréstimo?
 - abrir arquivos (*fechando outros ...*)
 - *carregar registros na memória (abre índice, usa ponteiro, estourou memória?,)*

Banco de dados

- Em uma fábrica com os dados em bancos de dados:

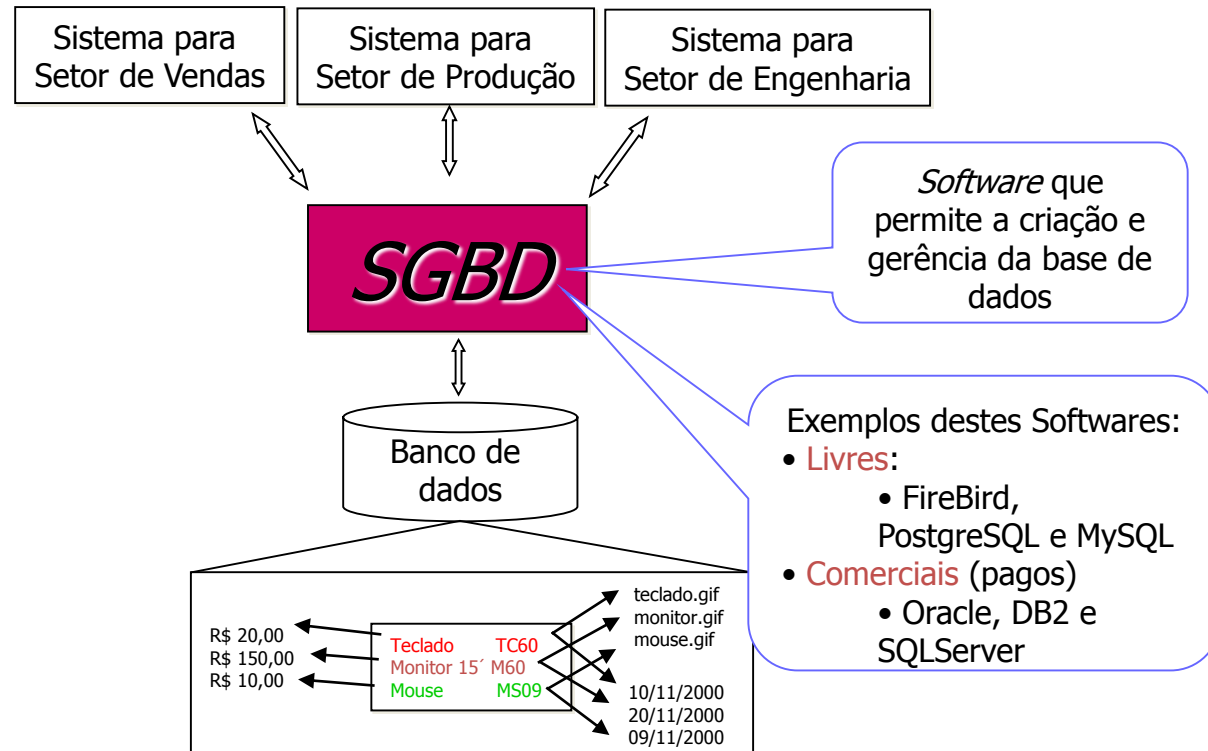


Dados aparecem
uma única vez
no banco

[baseado em Heuser]

Gerenciamento do banco de dados

■ BD de uma fábrica:



- Um SGBD (Sistema Gerenciador de Banco de Dados) consiste em uma coleção de dados inter-relacionados **e em um conjunto de programas para acessá-los**
- SGBDs são projetados para gerenciar grandes grupos de informações

- O gerenciamento envolve
 - A definição de estruturas para o armazenamento da informação
 - O fornecimento de mecanismos para manipular as informações
- Quando vários usuários acessam os dados o SGBD precisa garantir a INTEGRIDADE dos dados, evitando resultados anômalos

VANTAGENS DA UTILIZAÇÃO DA ABORDAGEM SGBD

- Controle de Redundância
- Restringindo Acesso Não Autorizado
- Garantindo o Armazenamento Persistente para Objetos Programas
- Garantindo o Armazenamento de Estruturas para o Processamento Eficiente de Consultas
- Garantindo Backup e Restauração
- Fornecendo Múltiplas Interfaces para os Usuários
- Representando Relacionamentos Complexos entre os Dados
- Forçando as Restrições de Integridade
- Permitindo Inferências e Ações Usando as Regras
- Implicações Adicionais do Uso da Abordagem de um Banco de Dados

- Essa redundância em armazenar os mesmos dados várias vezes gera muitos problemas.
 - Primeiro, há a necessidade de desempenhar uma única atualização lógica — como a entrada de dados de novos alunos — várias vezes: uma para cada arquivo no qual o dado do aluno estará armazenado. Isso gera uma duplicação de esforços.
 - Segundo, o espaço de armazenamento é desperdiçado quando o mesmo dado é armazenado repetidamente; esse problema pode ser sério para os bancos de dados grandes.
 - Terceiro, há a possibilidade de os arquivos que representam os mesmos dados se tornarem inconsistentes. Isso pode ocorrer porque uma atualização é aplicada somente a alguns arquivos, mas não em outros.

Autorizado

- Quando vários usuários utilizam um grande banco de dados, é provável que a maioria desses usuários não seja autorizada a acessar todas as informações disponíveis no banco de dados.
- A alguns usuários é permitido, apenas, consultar; outros podem consultar e atualizar os dados.
- O mais comum é fornecer aos usuários ou grupo de usuários contas protegidas por senhas, utilizadas para acessar o banco de dados.
- O SGBD deve garantir a segurança e um subsistema de autorização usado pelo DBA para criar contas

Garantindo o Armazenamento de Estruturas para o Processamento Eficiente de Consultas

- Os sistemas de banco de dados devem fornecer funcionalidades para a execução de atualizações e consultas eficientemente.
- Pelo fato de o banco de dados ser armazenado, tipicamente, em disco, o SGBD deve possuir estruturas de dados especializadas para aumentar a velocidade de pesquisa no disco dos registros desejados.
- Os arquivos auxiliares, chamados indexes (indexados), são utilizados com esse objetivo.
 - Os **indexes** são baseados em estruturas de dados árvores (tree) ou estruturas de dados hash, adequadamente adaptados para a pesquisa em disco.
- Os SGBD em geral têm um módulo de armazenamento temporário (buffering) que mantém partes do banco de dados armazenado na memória principal.
- Em outros casos, o SGBD pode utilizar o sistema operacional para fazer o armazenamento temporário dos dados no disco.
- O módulo do SGBD para o processamento de consulta e otimização é responsável pela escolha eficiente do plano de execução da consulta (query) baseado nas estruturas de armazenamento existentes. A opção de qual index criar e manter é parte do projeto físico do banco de dados e seu ajuste (tunning), que é de responsabilidade do DBA e sua equipe.
- 1

Garantindo Backup e Restauração

- Um SGBD deve prover facilidades para a restauração de falhas de hardware ou de software.
- O subsistema de backup e recuperação dos subsistemas do SGBD é responsável pela recuperação dessas falhas.
 - Por exemplo, se um sistema de computador falhar no meio de uma transação complexa de atualização, o subsistema de recuperação é responsável por garantir que o banco de dados seja recolocado no mesmo estado em que estava, antes do início da execução da transação.
 - Alternativamente, o subsistema pode assegurar que a transação seja resumida do ponto em que foi interrompida — sendo assim, seu efeito completo seria armazenado no banco de dados.

Fornecendo Múltiplas Interfaces para os Usuários

- Como diversos tipos de usuários com níveis de conhecimento técnico diferentes utilizam o banco de dados, o SGBD deve fornecer interfaces diferentes para esses usuários.
- Essas interfaces incluem linguagens de consulta para os usuários casuais; interfaces de linguagens de programação para programadores de aplicações; formulários e sequências de comandos para usuários parametrizáveis; interfaces de menus, interfaces de linguagem natural para usuários autônomos.
- Ambas, as interfaces com menus e aquelas com formulários, são comumente conhecidas como interfaces gráficas para os usuários — Graphical User Interfaces (GUIs).
- Muitos ambientes e linguagens especializadas existem para a especificação de GUIs.

Representando Relacionamentos Complexos entre os Dados

- Um banco de dados pode incluir uma grande variedade de dados que estão inter-relacionados de muitas maneiras
- O SGBD deve ter a capacidade de representar a variedade de relacionamentos complexos entre os dados, bem como recuperar e atualizar os dados relacionados fácil e eficientemente.

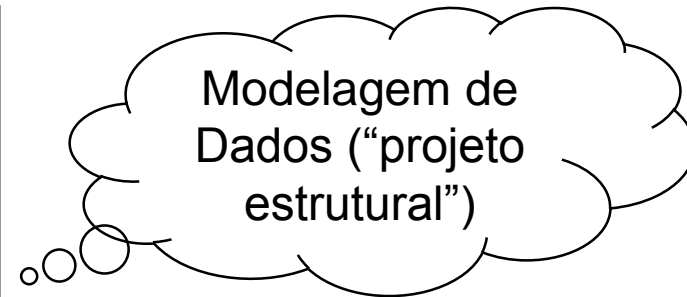
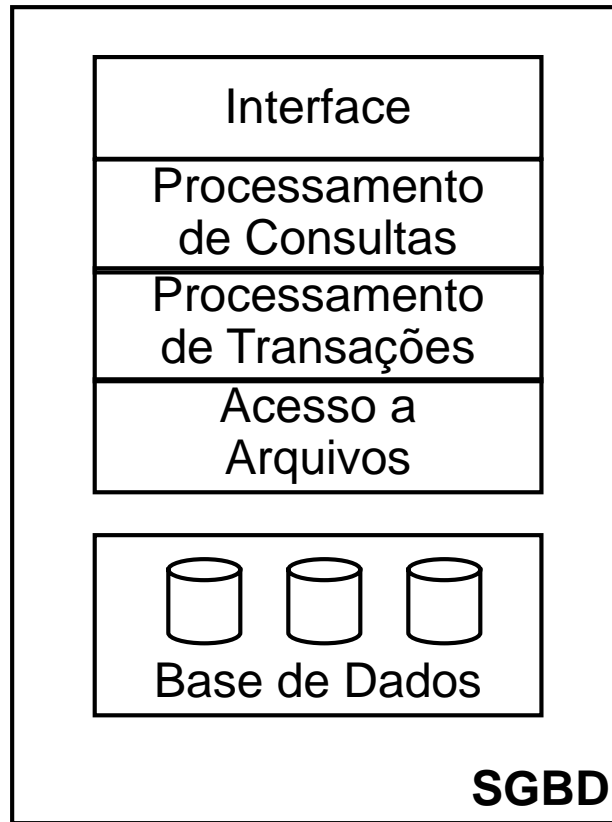
Forçando as Restrições de Integridade

- A maioria das aplicações de um banco de dados tem certas restrições de integridade que devem complementar os dados.
- O SGBD deve prover funcionalidades para a definição e a garantia dessas restrições.
- O tipo mais simples de restrição de integridade envolve a especificação de um tipo de dado para cada item de dados.
- Essas restrições são derivadas do significado ou da semântica dos dados e do minimundo que representam.
- Os projetistas do banco de dados são responsáveis por identificar as restrições de integridade durante o projeto do banco.
- Algumas restrições podem ser especificadas para o SGBD e executadas automaticamente.
- Outras podem ser testadas pelos programas de atualização ou no momento da entrada dos dados.

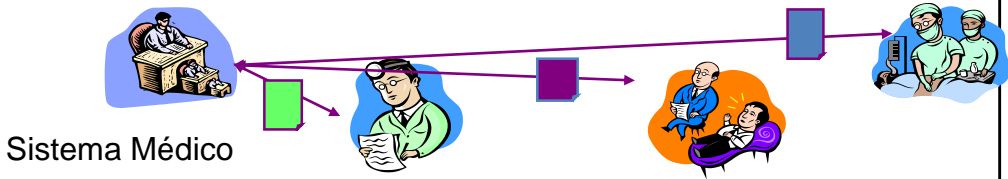
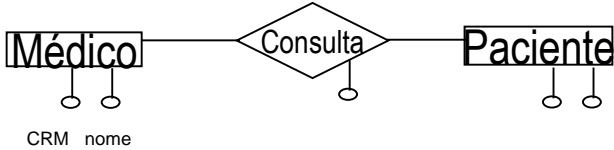
Usando as Regras

- Alguns sistemas de banco de dados oferecem capacidades para definir as regras de dedução por inferência gerando novas informações de fatos armazenados no banco de dados.
- Esses sistemas são chamados sistemas de banco de dados dedutivos.
 - Por exemplo, podem existir regras complexas no minimundo da aplicação para determinar quando um aluno está em recuperação. Isso pode ser especificado declaradamente como uma regra, e quando for compilada e mantida pelo SGBD,

Arquitetura Geral de um SGBD



Abstração de Dados

Mundo Real	 <p>Sistema Médico</p>			
Modelo Conceitual (modelo abstrato dos dados)	<ul style="list-style-type: none"> ♦ Independente do modelo de dados ♦ Independente do SGBD 			
Modelo Lógico (estrutura dos dados)	<ul style="list-style-type: none"> ♦ Dependente do modelo de dados ♦ Independente do SGBD <p>Médico (CRM, Nome)</p>	Relacional	Orientado a Objetos	Objeto-relacional
Modelo Físico	<ul style="list-style-type: none"> ♦ Dependente do modelo de dados ♦ Dependente do SGBD 	<ul style="list-style-type: none"> ♦ Organização física dos dados ♦ Estruturas de armazenamento de dados ♦ Índices de acesso 		

Modelos de Dados

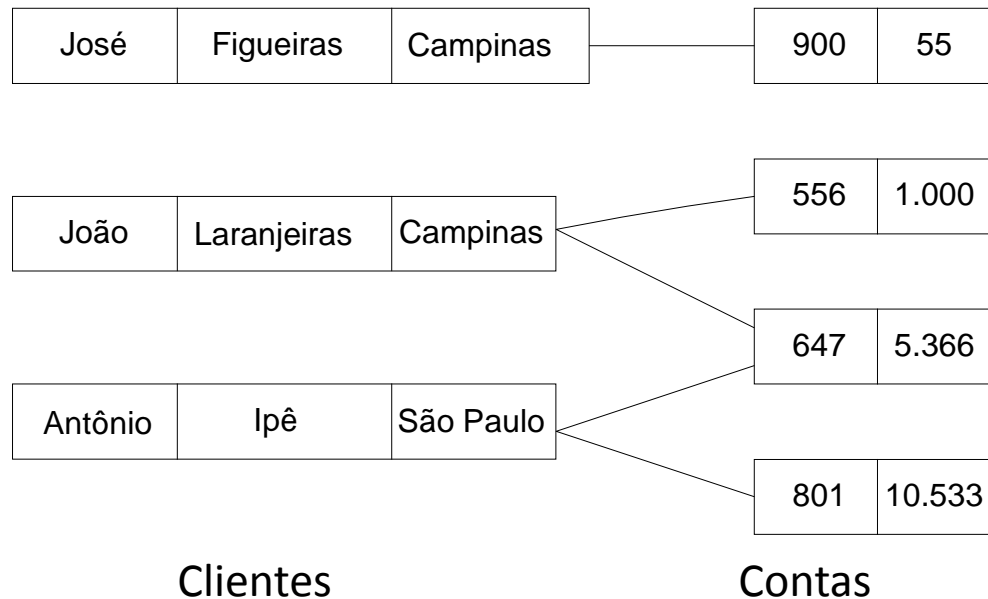
- Um **modelo de dados** é uma coleção de ferramentas conceituais para a **descrição** de dados, **relacionamentos**, semântica de dados e restrições de **consistência**

- Modelos de Dados (conceitual)
 - Entidade-Relacionamento (ER)
 - Orientado a Objetos (OO)
 - Modelos de Dados (lógicos)
 - Redes
 - Hierárquico
 - Relacional
 - Objeto-relacional
 - Orientado a Objetos
 - NoSql
- } Modelos mais antigos

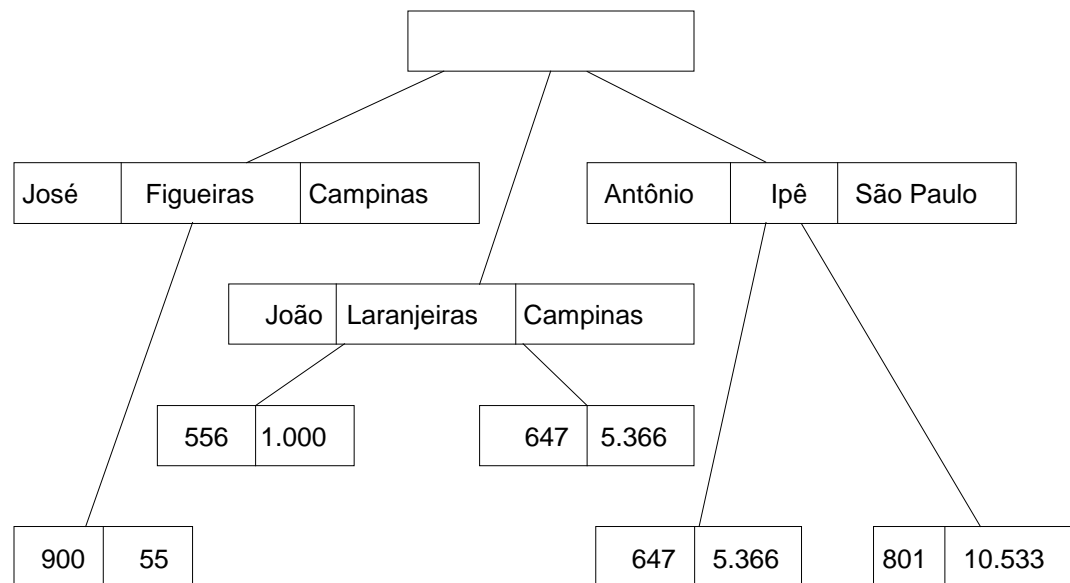
Exemplo das Informações em um Banco de Dados

nome	rua	cidade	conta	saldo
José	Figueiras	Campinas	900	55
João	Laranjeiras	Campinas	556	1.000
João	Laranjeiras	Campinas	647	5.366
Antônio	Ipê	São Paulo	647	5.366
Antônio	Ipê	São Paulo	801	10.533

Os dados são representados por coleções de registros e os relacionamentos por elos



- Os dados e relacionamentos são representados por registros e ligações, respectivamente.
- Os registros são organizados como coleções arbitrárias de árvores.



O Modelo Relacional

Tabela Cliente (dados)

cód-cliente	nome		
rua	cidade		
015 Campinas	José		Figueiras

Tabela Conta (dados)

nro-conta	
037 Paulo	João
900	Antônio

55
556
1.000
647
5.366
801

Tabela Cliente-Conta
(relacionamento)

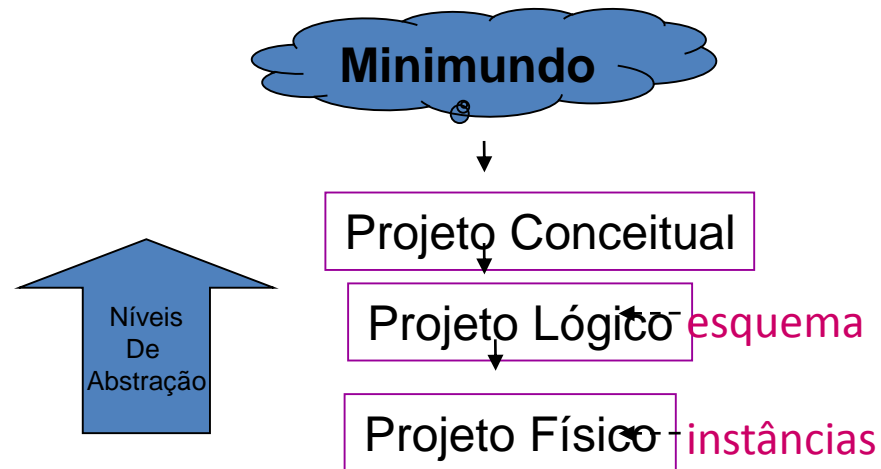
cód-cliente	nro-
conta	
015	900
021	556
021	647
037	647
037	801

- O modelo relacional não usa ponteiros ou ligações
- O modelo relacional relaciona registros a partir de valores do registro

- Os bancos de dados mudam a medida que informações são inseridas ou apagadas
 - A coleção de informações armazenadas é chamada de **INSTÂNCIA** do banco de dados (mudam com frequência)
 - O projeto geral do banco de dados é chamado **ESQUEMA** do banco de dados (não mudam com frequência)

- O uso de bancos de dados permite modificar o ESQUEMA dos dados em um nível sem afetar a definição do esquema em um nível mais alto. Isto é chamado de

***independência
dos dados***



- Existem 2 tipos de Independência
 - ***Independência física de dados:*** habilidade de modificar o **esquema físico** sem a necessidade de reescrever os programas aplicativos
 - Estas modificações são necessárias para melhorar o desempenho
 - ***Independência lógica de dados:*** habilidade de modificar o **esquema conceitual** sem a necessidade de reescrever os programas aplicativos
 - Estas modificações são necessárias quando a estrutura lógica é alterada.
 - Exemplo: adição de um novo atributo

- A independência lógica dos dados é mais difícil de ser alcançada do que a independência física, pois os programas são bastante dependentes da estrutura lógica dos dados que eles acessam

Linguagem de Definição de Dados (DDL)

- Um esquema de banco de dados é especificado por um conjunto de definições expressas por uma linguagem especial chamada ***linguagem de definição de dados (Data Definition Language)***
- *O resultado da compilação de comandos de uma DDL é o conjunto de tabelas que serão armazenadas no dicionário (ou diretório) de dados*

Linguagem de Definição de Dados (DDL)

- Um dicionário de dados contém metadados, i.e., dados sobre os dados
- Este dicionário (diretório) é consultado antes que os dados sejam lidos ou modificados no sistema de banco de dados

Linguagem de Manipulação de Dados (DML)

- Manipulação de dados significa:
 - A busca da informação armazenada no BD
 - A inserção de novas informações no BD
 - A eliminação de informações do BD
 - A modificação dos dados armazenados no BD
- No nível físico precisamos definir algoritmos que permitam acesso eficiente aos dados

Linguagem de Manipulação de Dados (DML)

- A linguagem de manipulação dos dados permite ao usuário manipular os dados da seguinte forma:
 - Procedural: o usuário informa qual dado deseja acessar e como obtê-lo
 - Não-procedural: o usuário informa qual dado deseja acessar SEM especificar como obtê-lo

Linguagem de Manipulação de Dados (DML)

- Linguagens não-procedurais são usualmente mais fáceis de aprender e usar do que DMLs procedurais
- Se o usuário NÃO especificar COMO obter os dados, as linguagens não-procedurais poderão gerar um código não tão eficiente.

Linguagem de Manipulação de Dados (DML)

- Uma consulta (QUERY) é um comando de busca de uma informação no BD
- A parte da DML que busca informações é chamada **LINGUAGEM DE CONSULTA**

- Programadores de Aplicativos:
 - São os usuários que escrevem os programas de aplicação através da DML
 - Exemplos de um sistema bancário são programas que geram cheques, fazem débitos e créditos em contas, transferem fundos entre contas
- Usuários de alto nível
 - Interagem com o sistema sem escrever programas
 - Formulam consultas em uma linguagem de consulta, e cada consulta é submetida a um processador de consulta, cuja função é gerar um comando da DML

- Usuários especializados (especialistas)
 - Escrevem aplicativos especializados como sistemas especialistas
- Usuários ingênuos
 - Interagem com o sistema invocando os programas aplicativos
 - Exemplo: um cliente do banco invocaria um programa para efetuar a transferência de 50 reais da conta A para a conta B

- **Administrador do banco de dados:** tem o controle central dos dados e dos programas de acesso aos dados

- **Funções do Administrador do banco de dados:**
 - Definição do esquema
 - Definição de estruturas de armazenamento e métodos de acesso
 - Modificação de esquema e de organização física
 - Concessão de autorização para acesso aos dados
 - Especificação de restrições de integridade

- Gerenciador de arquivos
- Gerenciador do banco de dados
- Processador de consultas
- Pré-compilador da DML
- Compilador da DDL
 - Arquivos de dados
 - Dicionário de dados
 - Índices

- Utilizadas para controlar a integridade dos dados no Banco de dados
 - Acessos simultâneos vários usuários
 - Falhas no sistema

- Escolhe a forma mais eficiente para execução de uma consulta

Arquitetura Geral de um SGBD

