

Fundamentos de Banco de Dados

# Aula 1 – 01

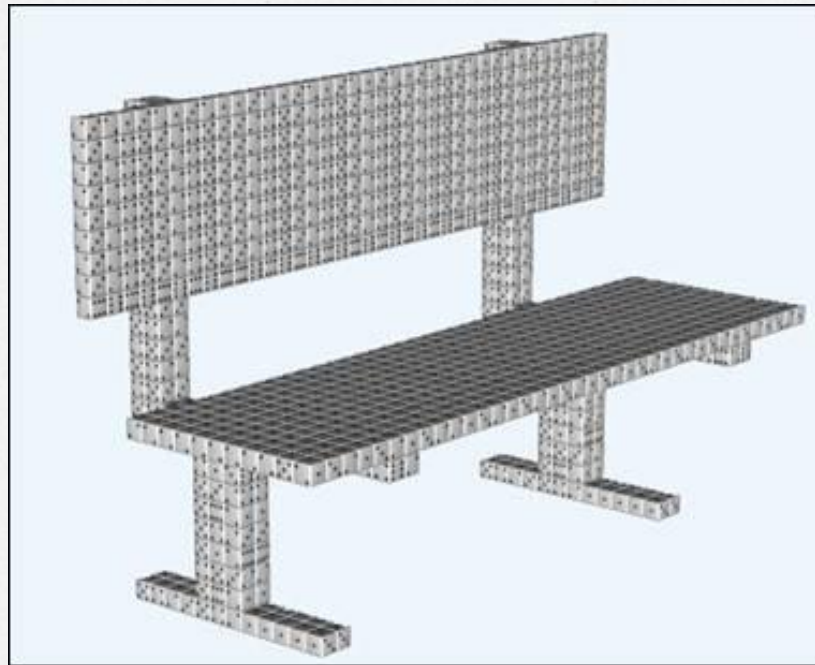
# Apresentação

- o Fale um pouco sobre você:
  - o Nome:
  - o Semestre:
  - o Experiência Profissional:
  - o Empresa que trabalha:
  - o Expectativa de aprendizagem na disciplina:



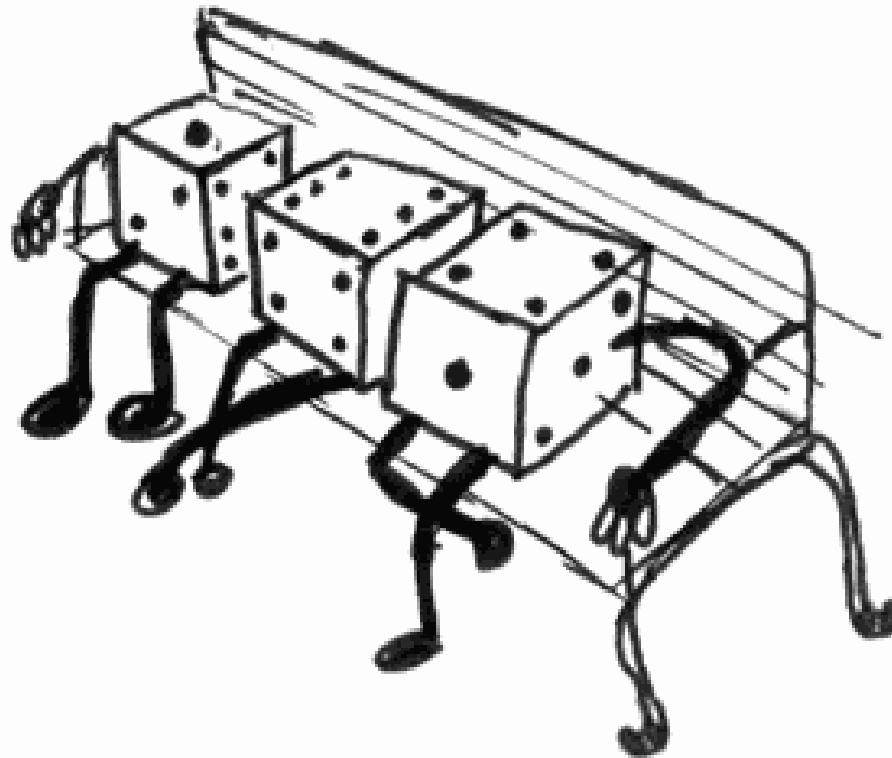
# Conceitos Iniciais

o Banco de Dados!



# Conceitos Iniciais

O BANCO DE DADOS



# Conceitos Iniciais

- o O que é Banco de Dados?
- o Para que serve?
- o Onde encontramos?
- o Sua utilização é adequada, eficiente?



# Conceitos Iniciais

- o Banco de Dados: é uma coleção de dados inter relacionados, representando informações sobre um domínio específico.
- o Dados são considerados fatos conhecidos que podem ser registrados e possuem significado implícito.
- o Exemplos: lista telefônica, controle do acervo de uma biblioteca, sistema de controle dos recursos humanos de uma empresa.

# Conceitos Iniciais

- o Sistema de Gerenciamento de Bancos de Dados (SGBD):
  - o É um software com recursos específicos para facilitar a manipulação das informações dos bancos de dados e o desenvolvimento de programas aplicativos.
- o Exemplos: Oracle, SQL Server, My SQL, PostgreSQL, Paradox, Access



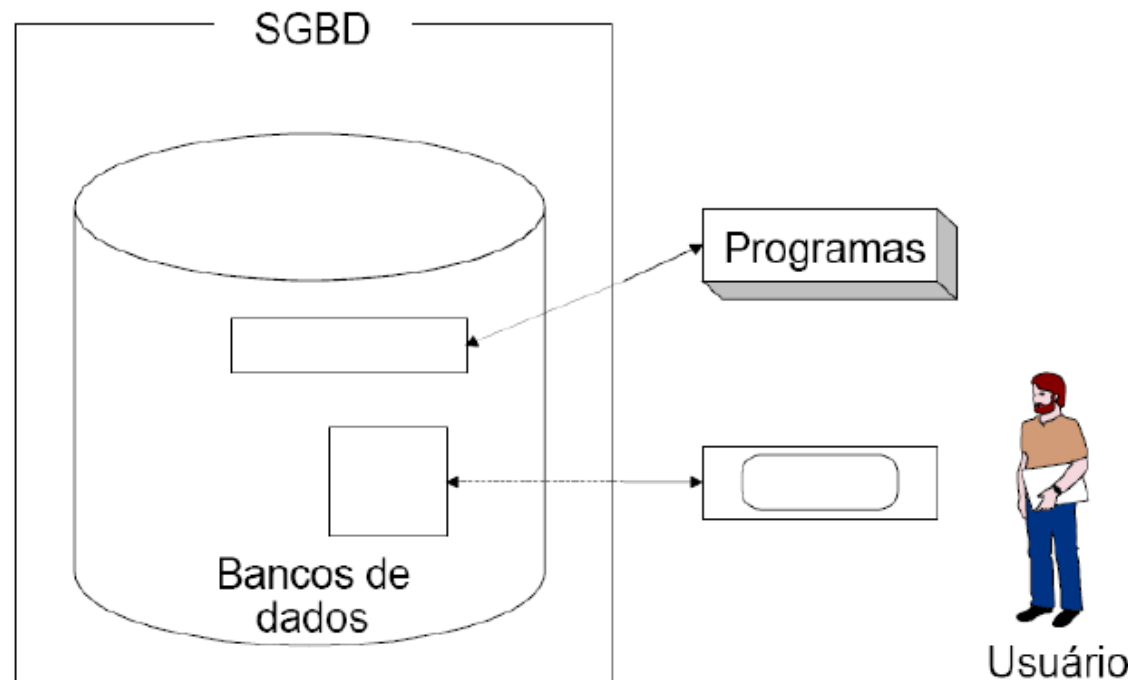
# Conceitos Iniciais

- o Sistema de Bancos de Dados
  - o É um sistema de manutenção de registros por computador, envolvendo quatro componentes principais:
    - o dados,
    - o hardware,
    - o software e
    - o usuários.
  - o O sistema de bancos de dados pode ser considerado como uma sala de arquivos eletrônica.
  - o Existe uma série de métodos, técnicas e ferramentas que visam sistematizar o desenvolvimento de sistemas de bancos de dados.

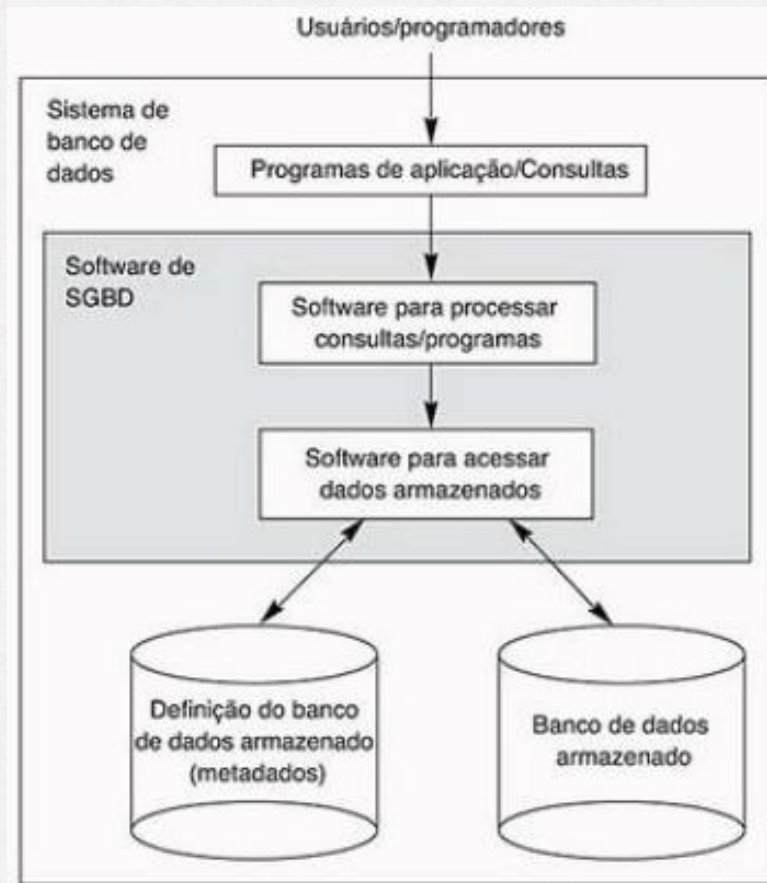


# Conceitos Iniciais

## Os Componentes de um Sistema de Bancos de Dados



# Conceitos Iniciais





# Conceitos Iniciais

- o Objetivos de um sistema de Banco de Dados
  - o Isolar os usuários dos detalhes mais internos do banco de dados (abstração de dados).
  - o Prover independência de dados às aplicações (estrutura física de armazenamento e à estratégia de acesso).

# Conceitos Iniciais

## o Vantagens

- o rapidez na manipulação e no acesso à informação,
- o redução do esforço humano (desenvolvimento e utilização),
- o disponibilização da informação no tempo necessário,
- o controle integrado de informações distribuídas fisicamente,
- o redução de redundância e de inconsistência de informações,
- o compartilhamento de dados,
- o aplicação automática de restrições de segurança,
- o redução de problemas de integridade.



# Conceitos Iniciais

## o Desvantagens

- o Sem dispositivos de controle adequados, a segurança pode ficar comprometida; por exemplo, no caso de acesso não autorizado a dados.
- o A integridade das informações pode ser comprometida se não houver mecanismos de controle; por exemplo no caso de manipulação concorrente de dados.
- o A operação do sistema de banco de dados e o desenvolvimento de aplicações precisam ser feitos com muita precisão para evitar que informações não correspondam à realidade.
- o A administração do sistema de banco de dados pode se tornar muito complexa em ambientes distribuídos, com grande volume de informações manipuladas por uma grande quantidade de usuários.

# Conceitos Iniciais

## o Modelo Relacional:

- o é um modelo lógico de dados, baseado em registros, que provê abstração de dados nos níveis conceitual e de visões do usuário. Neste modelo, tabelas representam dados e relacionamentos entre dados.



# Conceitos Iniciais

- o **Linguagem de definição de dados (DDL):**
  - o define a estrutura (esquema) de um banco de dados, principalmente nos níveis conceitual e de visões de usuário. A compilação de definições em DDL é armazenada no dicionário de dados que, portanto, encerra dados sobre dados (metadados).

# Conceitos Iniciais

- o **Linguagem de manipulação de dados (DML)**
  - o permite aos usuários e aplicações acessar ou manipular as informações contidas num banco de dados. A manipulação de dados engloba recuperação, inserção, exclusão e modificação da informação armazenada.



# Conceitos Iniciais

## o Administrador de Banco de Dados:

- o definição e atualização do esquema dos bancos de dados.
- o definição da estrutura de armazenamento e da estratégia (ou método) de acesso;
- o concessão de autorização para acesso a dados;
- o definição de controles de integridade;
- o definição de estratégias de cópia de segurança e recuperação;
- o monitoração de desempenho;
- o execução de rotinas de desempenho;
- o atualização da organização física.

# Histórico

- o O primeiro Sistema Gerenciador de Banco de Dados (SGBD) comercial surgiu no final de **1960** com base nos primitivos sistemas de arquivos disponíveis na época, os quais **não controlavam o acesso concorrente por vários usuários ou processos**.
- o Os SGBDs **evoluíram** desses sistemas de arquivos de armazenamento em disco, criando **novas estruturas de dados** com o objetivo de armazenar informações. Com o tempo, os SGBD's passaram a utilizar diferentes formas de representação, ou modelos de dados, para descrever a estrutura das informações contidas em seus bancos de dados.
- o Atualmente, os seguintes modelos de dados são normalmente utilizados pelos SGBD's: **modelo hierárquico, modelo em redes, modelo relacional (amplamente usado) e o modelo orientado a objetos**.



# Modelo de Dados

- o Modelo Hierárquico
  - o O modelo hierárquico foi o primeiro a ser reconhecido como um modelo de dados.
  - o Nesse modelo de dados, são estruturados em hierarquias ou árvores.
  - o Os nós das hierarquias contêm ocorrências de registros, onde cada registro é uma coleção de campos (**atributos**), cada um contendo apenas uma informação.
  - o O registro da hierarquia que precede a outros é o registro-pai, os outros são chamados de registros-filhos.

# Modelo de Dados

- o Modelo em Rede

- o O modelo em redes surgiu como uma extensão ao modelo hierárquico, eliminando o conceito de hierarquia e permitindo que um mesmo registro estivesse envolvido em várias associações.
- o No modelo em rede, os registros são organizados em grafos onde aparece um único tipo de associação



# Modelo de Dados

## o Modelo Relacional

- o O modelo relacional apareceu devido às seguintes necessidades:
  - o *aumentar a independência de dados* nos sistemas gerenciadores de banco de dados;
  - o *prover um conjunto de funções* apoiadas em álgebra relacional para armazenamento e recuperação de dados;
  - o *permitir processamento ad hoc*
- o O modelo relacional, tendo por base a teoria dos conjuntos e álgebra relacional, foi resultado de um estudo teórico realizado por CODD

# Modelo de Dados

## o Modelo Relacional

- o O Modelo relacional revelou-se ser o mais flexível e adequado ao solucionar os vários problemas que se colocam no nível da concepção e implementação da base de dados.
- o A estrutura fundamental do modelo relacional é a relação (tabela). Uma relação é constituída por um ou mais atributos (campos) que traduzem o tipo de dados a armazenar.
- o Cada instância do esquema (linha) é chamada de tupla (registro).
- o O modelo relacional não tem caminhos pré-definidos para se fazer acesso aos dados como nos modelos que o precederam.



# Modelo de Dados

- o Modelo Orientado a Objetos
  - o A motivação para seu surgimento está em função dos limites de armazenamento e representação semântica impostas no modelo relacional.
  - o Alguns exemplos são os sistemas de informações geográficas (SIG).

# Modelo de Dados

- o Sistemas Objeto-Relacionais
  - o A área de atuação dos sistemas Objeto-Relacional tenta suprir a dificuldade dos sistemas relacionais convencionais, que é o de representar e manipular dados complexos, visando ser mais representativos em semântica e construções de modelagens.
  - o A solução proposta é a adição de facilidades para manusear tais dados utilizando-se das facilidades SQL (Structured Query Language) existentes.
  - o Para isso, foi necessário adicionar: extensões dos tipos básicos no contexto SQL; representações para objetos complexos no contexto SQL; herança no contexto SQL e sistema para produção de regras.



# Arquiteturas de SGBDs

- o Plataformas centralizadas.
  - o Na arquitetura centralizada, existe um computador com grande capacidade de processamento, o qual é o hospedeiro do SGBD e emuladores para os vários aplicativos.
  - o Esta arquitetura tem como principal vantagem a de permitir que muitos usuários manipulem grande volume de dados.
  - o Sua principal desvantagem está no seu alto custo, pois exige ambiente especial para mainframes e soluções centralizadas.

# Arquiteturas de SGBDs

- o Sistemas de Computador Pessoal - PC.
  - o Os computadores pessoais trabalham em sistema stand-alone, ou seja, fazem seus processamentos sozinhos.
  - o No começo esse processamento era bastante limitado, porém, com a evolução do hardware, tem-se hoje PCs com grande capacidade de processamento.
  - o Eles utilizam o padrão Xbase e quando se trata de SGBDs, funcionam como hospedeiros e terminais.
  - o Desta maneira, possuem um único aplicativo a ser executado na máquina. A principal vantagem desta arquitetura é a simplicidade.



# Arquiteturas de SGBDs

- o Banco de Dados Cliente-Servidor.
  - o Na arquitetura Cliente-Servidor, o cliente (front\_end) executa as tarefas do aplicativo, ou seja, fornece a interface do usuário (tela, e processamento de entrada e saída).
  - o O servidor (back\_end) executa as consultas no DBMS e retorna os resultados ao cliente.
  - o Apesar de ser uma arquitetura bastante popular, são necessárias soluções sofisticadas de software que possibilitem: o tratamento de transações, as confirmações de transações (commits), desfazer transações (rollbacks), linguagens de consultas (stored procedures) e gatilhos (triggers).
  - o A principal vantagem desta arquitetura é a divisão do processamento entre dois sistemas, o que reduz o tráfego de dados na rede.

# Arquiteturas de SGBDs

- o Banco de Dados Distribuídos (N camadas).
  - o Nesta arquitetura, a informação está distribuída em diversos servidores. Cada servidor atua como no sistema cliente-servidor, porém as consultas oriundas dos aplicativos são feitas para qualquer servidor indistintamente.
  - o Caso a informação solicitada seja mantida por outro servidor ou servidores, o sistema encarrega-se de obter a informação necessária, de maneira transparente para o aplicativo, que passa a atuar consultando a rede, independente de conhecer seus servidores.
  - o Exemplos típicos são as bases de dados corporativas, em que o volume de informação é muito grande e, por isso, deve ser distribuído em diversos servidores.
  - o Porém, não é dependente de aspectos lógicos de carga de acesso aos dados, ou base de dados fracamente acopladas, em que uma informação solicitada vai sendo coletada numa propagação da consulta numa cadeia de servidores.
  - o A característica básica é a existência de diversos programas aplicativos consultando a rede para acessar os dados necessários, porém, sem o conhecimento explícito de quais servidores dispõem desses dados.



# Fatores de Influência na escolha de um BD

## o Controle de redundância:

- o O banco de dados deve ser capaz de garantir que os dados não tenham duplicidade.
- o Isso normalmente é conhecido como integridade referencial.
- o Desta forma, não seria possível incluir dois registros com o mesmo código (chave primária).
- o Também não seria possível excluir um registro que tivesse relacionamento com outras tabelas (chave estrangeira).
- o Esta integridade é a base do modelo relacional, portanto é necessário que o banco de dados tenha a capacidade de gerenciar o controle de redundância.

## Fatores de Influência na escolha de um BD

### o **Compartilhamento de dados:**

- o a informação deve estar disponível para qualquer número de usuários de maneira rápida, concomitante e segura.
- o É impensável, nos dias atuais, imaginar um banco de dados exclusivo para um usuário.
- o A informação, cada vez mais, deve ser compartilhada por diversas pessoas da empresa.
- o Disponibilizar a informação com rapidez e segurança é requisito fundamental para determinar a escolha do banco de dados.



## Fatores de Influência na escolha de um BD

### o Controle de acesso:

- o é essencial saber quem fez e o que cada usuário pode fazer dentro do banco de dados.
- o Disponibilizar a informação é pouco.
- o Deve haver controle sobre o que é disponibilizado.
- o Deve-se analisar as possibilidades de controle de acesso às tabelas e colunas do banco de dados e às operações que cada usuário pode realizar (inclusão, alteração, consulta ou exclusão).

## Fatores de Influência na escolha de um BD

### o Cópias de Segurança:

- o Deve haver rotinas específicas para realizar cópias de segurança dos dados armazenados.

### o Versões Express?



# Fatores de Influência na escolha de um BD

## o Suporte às Transações:

- o as transações são originadas em qualquer operação que seja feita nos dados armazenados.
- o Realizar o controle sobre essas transações, garantindo a integridade das informações armazenadas mesmo quando há diversos usuários realizando operações ao mesmo tempo, é uma necessidade cada vez mais importante para os bancos de dados.
- o Há diversos níveis para o controle de transações. O mínimo necessário para os dias atuais é o bloqueio por linha, ou seja, cada alteração bloqueará apenas uma linha no banco de dados.
- o Com isso, há uma maior disponibilidade da informação armazenada visto que poucas linhas estarão efetivamente bloqueadas por transações pendentes.

# Fatores de Influência na escolha de um BD

## o Suporte a Programação:

- o mesmo para quem utiliza arquitetura de desenvolvimento em três ou mais camadas, algumas operações continuam sendo mais rápidas se forem realizadas diretamente no banco de dados.
- o Com isso, o banco de dados deve possuir uma linguagem de programação que permita realizar rotinas específicas diretamente sobre ele.
- o Além disso, muitas regras de negócio são implementadas diretamente no banco de dados.
- o Por exemplo, realizar o pedido de compra toda vez que o estoque do produto chegar ao nível mínimo ou bloquear a venda, caso o cliente esteja comprando além do seu limite de crédito.



# Fatores de Influência na escolha de um BD

## o Recuperação:

- o Falhas acontecem.
- o O que fazer quando houver uma quebra total do banco de dados?
- o Qual o caminho para recuperação deste desastre?
- o Claro que um bom backup pode resolver boa parte dos problemas, mas o que, além disso, o banco de dados poderá fazer?
- o Mecanismos de backup online e em diversos servidores e clusters, entre outros, são ferramentas importantes quando um problema acontece.

## Fatores de Influência na escolha de um BD

### o Desempenho:

- o Desempenho: de nada adianta ter um banco de dados completo se este for lento para as necessidades da empresa.
- o O desempenho do banco de dados muitas vezes pode ser melhorado com técnicas de tuning (ajuste) realizadas diretamente no banco.
- o Convém ter certeza de que o banco de dados permite realizar estes ajustes, se eles podem ser realizados e em qual escala.



## Fatores de Influência na escolha de um BD

### o Escalabilidade:

- o É necessário saber os limites do banco de dados.
- o Convém, principalmente para os bancos de dados livres, considerar o tamanho máximo do banco de dados e o número máximo de linhas em cada tabela.
- o Conhecer casos de sucesso é fundamental para determinar se o banco de dados está dentro da necessidade da empresa.

# Banco de Dados Relacionais

## Objetivos e Requisitos de Um Banco de Dados:

- o Absorver a expansão das atividades da empresa, possibilitando sistemas gerenciais cada vez mais complexos;
- o Permitir maior facilidade no desenvolvimento de sistemas;
- o Prover mecanismos de controle centralizado das informações e representação de dados mais próxima da realidade;
- o Eleger bancos de dados que comportem as últimas inovações do mercado ==> vários tipos de dados: som, imagem em movimento, imagem tridimensional, gráficos, textos, e suas combinações;
- o Assegurar a facilidade de manutenção de dados e instrumentos de rápida restauração em situações de erro;



# Banco de Dados Relacionais

## o Principais Características:

- o Estrutura tabular: os dados são representados na forma de tabelas, nas quais temos tuplas (linhas) e atributos ou domínios (colunas).
- o Álgebra Relacional: a manipulação destas tabelas é feita através de operadores que permitem qualquer tipo de acesso a uma tabela ou a um conjunto de tabelas (seleção, união, junção - join etc)
- o Dicionário de Dados: repositório onde são armazenados os dados. Deve ser integro, ativo e dinâmico.

# Banco de Dados Relacionais

## o Principais Características:

- o Esquema Conceitual: representa o ponto de vista global da empresa sobre a organização dos dados ==> AD (Administrador Dados): identifica padrão de entidades, atributos e relacionamentos, determina restrições de integridade e segurança ao modelo de dados.
- o Esquema Externo: representa o ponto de vista particular de cada usuário sobre a organização dos dados ==> AA (Administrador de Aplicação): agrupar informações e definir critérios de acesso/atualização;
- o Esquema Físico: tem a ver com o armazenamento e a manipulação da informação. DBA (Administrador da Base de Dados): mapeamento do modelo lógico para o físico - definir estruturas de acesso, índices, modo de transmissão de dados ...



# Banco de Dados Relacionais

## Regras para um BD Relacional:

- Gerenciar o Banco de Dados exclusivamente através de capacidades relacionais;
- Toda a informação é apresentada sempre em nível lógico por valores em tabelas;
- Todo o dado tem a garantia de ser logicamente acessível;
- Valores nulos representados por ausência de valores (não um valor zero ou caracter especial);
- Catálogo baseado no modelo relacional;
- Deve haver pelo menos uma linguagem que permita: definição dos dados, definição de visão, manipulação dos dados, restrições de integridade, autorização, limites de transação (commit, roolback);

# Banco de Dados Relacionais

## o Regras para um BD Relacional:

- o Todas as visões são atualizáveis;
- o Inserção, atualização e remoção de alto nível;
- o Independência de dados físicos: : programas permanecem inalterados quando ocorre mudança nos métodos de acesso ou nas representações de memória;
- o Independência de dados lógicos: programas permanecem inalterados quando ocorre mudança na BD;
- o Independência de integridade: definida na linguagem relacional;
- o Independência de distribuição: não ocorre alteração nas aplicações quando há redistribuição dos dados ou na distribuição inicial;
- o Não subversão: linguagem de acesso não pode ferir princípios de integridade;



# Metodologia de Desenvolvimento para Banco de Dados Relacional:

## o **Estratégia: análise das necessidades;**

- o Análise: modelo conceitual do BD
- o Projeto: modelo lógico do BD
- o Construção: projeto físico do BD
- o Implementação: Instalação do BD
- o Implantação: monitoração
- o Implantação: sintonização

# Metodologia de Desenvolvimento para Banco de Dados Relacional:

- o **Estratégia:==> obter uma especificação das necessidades globais**
  - o levantamento dos Planos Empresariais;
  - o levantamento das necessidades gerais de informação;
  - o administração de dados;
  - o necessidade de processamento;
  - o estudo de viabilidade;
  - o utilização do dicionário de dados;



# Metodologia de Desenvolvimento para Banco de Dados Relacional:

- o **Análise: ==> gerar esquema conceitual do BD**
  - o estruturação das necessidades globais de informação;
  - o desenho do Diagrama Hierárquico de Funções;
  - o estruturação opcional do Diagrama de Contexto;
  - o desenho opcional do Diagrama de Fluxo de Dados;
  - o especificação das entidades e relacionamentos;
  - o geração do diagrama de entidade/relacionamento;
  - o definição de eventos, gatilhos e procedimentos;
  - o balanceamento entre entidades e funções (mapeamento);

# Metodologia de Desenvolvimento para Banco de Dados Relacional:

- o **Projeto: ==> cria estrutura lógica do BD - independente de SGBD**
  - o análise das opções de montagem do BD;
  - o normalização das entidades, atributos e seus relacionamentos;
  - o escolha das chaves candidatas que se transformam em chaves primárias;
  - o consolidação das entidades com mesma chave primária, adequando conceitos de super e subtipo;
  - o mapeamento do esquema conceitual do BD (ER)



# Metodologia de Desenvolvimento para Banco de Dados Relacional:

- o **Construção: ==> transformar o modelo lógico no formato adequado a um SGBD específico**
  - o análise das opções de computação e sistema operacional;
  - o especificação das características físicas do BD;
  - o criação de estruturas (clusters) para melhorar desempenho;
  - o possível “desnormalização” visando maior rapidez nas consultas;
  - o mapeamento das consultas no BD;
  - o desenho das telas para os usuários;

# Metodologia de Desenvolvimento para Banco de Dados Relacional:

## o Implementação: ==> instalação do BD

- o geração e processamento das definições do BD na linguagem de definições do SGBD (ex.: SQL);
- o geração e replicação das definições do BD em cada servidor da rede;
- o construção dos programas;
- o carga inicial do Banco de Dados;
- o geração da segurança de acesso ao BD;



# Metodologia de Desenvolvimento para Banco de Dados Relacional:

- o **Implantação - Monitoração: ==> medir desempenho**
  - o acompanhamento da utilização do BD;
  - o análise da evolução do desempenho;

# Metodologia de Desenvolvimento para Banco de Dados Relacional:

- o **Implantação - Sintonização (Tunning): ==> crítica sólida do desempenho**
  - o análise de pontos críticos de desempenho;
  - o especificação de novas estruturas;
  - o otimização do BD central e nos vários servidores



## Modelos de Banco de Dados Relacionais:

- o DB2
- o Ingres
- o InterBase
- o MySQL
- o Oracle
- o PostgreSQL
- o Progress
- o Microsoft SQL Server
- o Sybase
- o Informix
- o Firebird