



MINISTÉRIO DA CIÊNCIA E TECNOLOGIA  
**INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS**

# Bancos de Dados Geográficos - 2010

Lúbia Vinhas  
Gilberto Câmara



# Introdução

- Essa disciplina visa dar aos alunos dos cursos CAP e SERE a compreensão sobre conceitos de Sistemas Gerenciadores de Bancos de Dados e sua aplicação ao domínio da Geoinformação
- Propomos um enfoque híbrido: não é um curso forte de SGBD para alunos da Computação, nem um curso de somente de aplicação para alunos do Sensoriamento Remoto
- Espera-se que os alunos do Sensoriamento Remoto aprendam a modelar corretamente seus problemas, possam avaliar diferentes sistemas e fazer o melhor uso deles
- Espera-se que os alunos do Computação aprendam um uso não convencional de SGBDs relacionais e objeto-relacionais, de modo a serem capazes de produzir ferramentas mais adequadas ao domínio da geoinformação



# Dinâmica do curso

- Os docentes da disciplina são os professores Gilberto Câmara ([gilberto.camara@inpe.br](mailto:gilberto.camara@inpe.br)) e Lúbia Vinhas ([lubia@dpi.inpe.br](mailto:lubia@dpi.inpe.br))
  - Enviar um e-mail para Lúbia da conta de correio que você acompanha. Qualquer aviso será enviado por correio eletrônico.
- As aulas são as segundas e sextas, das 10:00 as 12:00, na sala 2 da rotunda. Se houver necessidade de mudança de horário ou local, avisarei com antecedência por e-mail até o dia anterior.
- Quando possível, aulas práticas serão agendadas para o mesmo horário, usando o Lab Geo (prédio novo ao lado do SERE) . Essas aulas serão confirmadas ao longo do curso.
- Página da disciplina: <http://wiki.dpi.inpe.br/doku.php?id=cap349>



# Avaliação

- A nota final do curso será dada pela seguinte fórmula:

$$NF = NL * 0.2 + NP * 0.3 + NT * 0.5$$

Onde:

NF = Notal final

NL = Nota de listas (individual ou em duplas)

NP = Nota da prova (individual)

NT = Nota do trabalho proposto pelo ALUNO (individual)



# Referências

1. Shekhar, S. and S. Chawla (2003). Spatial databases - a tour. Upper Saddle River, NJ, USA, Prentice-Hall
  2. Rigaux, P., M. Scholl, et al. (2002). Spatial Databases with Application to GIS. San Francisco, Morgan Kaufman
  3. Casanova, M., G. Câmara, et al. (2005). Bancos de Dados Geográficos. Curitiba
- 
- **Notas de Aulas:** são baseadas no material dos anos anteriores criados pelo Gilberto Ribeiro e Karine Reis, material de ensino dos livros 1 e 2 acima e diversos outros materiais disponíveis na internet.



MINISTÉRIO DA CIÊNCIA E TECNOLOGIA  
**INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS**

# Motivação



# Dados Espaciais

- Dados Espaciais
  - Dados que contém a uma localização espacial em algum sistema de referência. Informação sobre a localização e forma (representadas por coordenadas) e as relações (topológicas) entre feições geográficas
- Sistemas de Referência
  - São basicamente as diferentes perspectivas de um observador a descrição de medidas (ex. posição) e sistemas de coordenadas são as diferentes formas de descrever medidas sob essas perspectivas
- Dados Geográficos ou Geoespaciais
  - São aqueles onde o sistema de referência é a superfície da terra



# Dados Espaciais

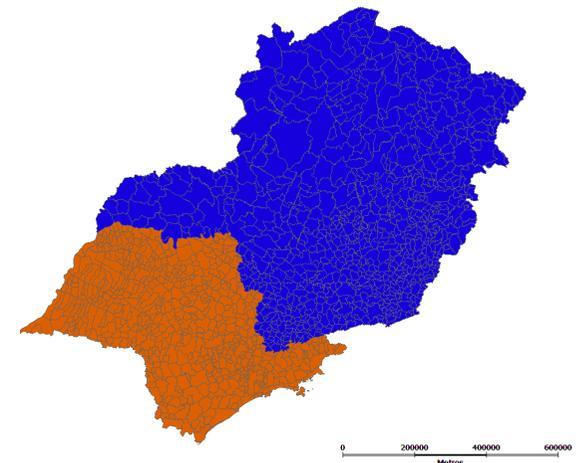
A nível conceitual, um objeto geográfico corresponde a um entidade do mundo real que possui duas componentes:

1. um conjunto de atributos alfa-numéricos, ou atributos descritivos
2. uma componente espacial, que pode conter a geometria (localização e forma) e a topologia (relacionamentos espaciais com outros objetos)

Objetos geográficos podem ser atômicos ou complexos

Como são obtidos os dados espaciais:

- através de levantamentos de campo (ex. GPS) e através de sensores remotos
- digitalizados a partir de mapas já existentes





# O que há de especial sobre dados espaciais?

- Os sistemas gerenciadores de banco de dados tradicionais (SGBDs) são muito bons em gerenciar dados não espaciais como strings, números e datas. Eles oferecem:
  - acesso eficiente aos dados através de linguagens de consulta declarativas
  - controle de transação e concorrência
  - controle de acesso e mecanismos de segurança
  - independência entre o modelo lógico e o modelo físico dos dados
- Porque não usar um SGBD tradicional para gerenciar dados espaciais?
- Dados espaciais são mais complexos. Alguns fatores complicantes:
  - modelagem de geometrias
  - necessidade de processar consultas complexas sobre relacionamentos espaciais
  - inadequação da estruturas de indexação tradicionais (uni-dimensionais)



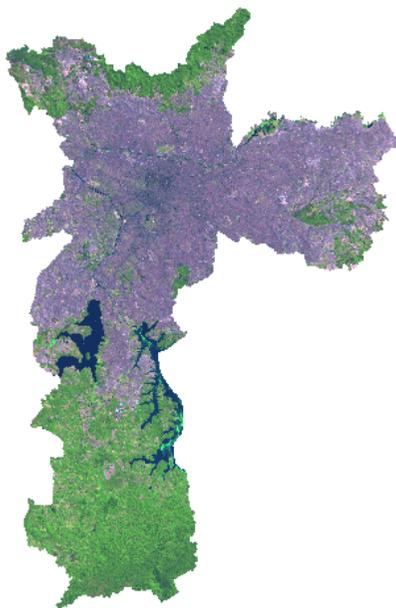
# Sistemas de Informação Geográfica

- **Consequência:** técnicas e extensões aos SGBDs (relacionais) tradicionais são necessárias para atender aos requisitos da gerência de dados espaciais
- Um SIG é um software que permite o armazenar, gerenciar e visualizar dados espaciais e fornece ferramentas de análise como:
  - **Consulta:** por região, por coordenadas, classificação
  - **MNT:** declividade, rede de drenagem, bacias
  - **Rede:** conectividade, caminho mínimo
  - **Distribuição:** detecção de mudanças, proximidade
  - **Análise/Estatística espacial:** padrões, auto-correlação
  - **Medidas:** distância, forma, adjacência, direção, perímetro

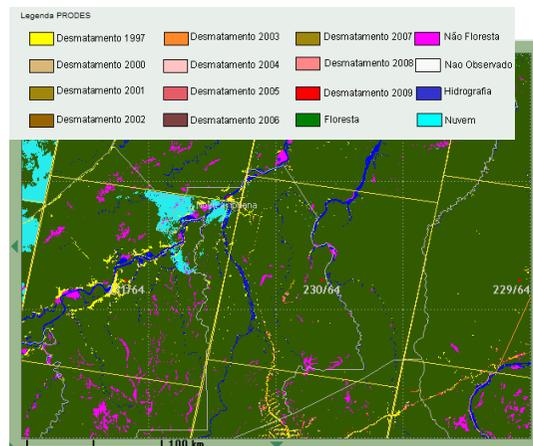


# Dados Espaciais

- Os SIG's são capazes de manipular diferentes tipos de dados espaciais



Imagens SR

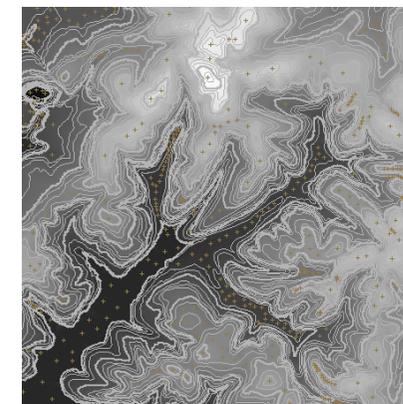


Dados Temáticos

Fotos aéreas



Topografia



SIG não trata apenas de mapas, mas a informação espacial é normalmente representada em mapas

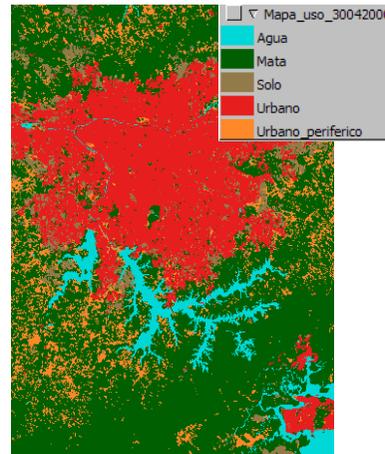


# Aplicações dos SIGs

- Interesse sobre dados espaciais e SIG está aumentando dramaticamente para além das aplicações tradicionais que incluem:



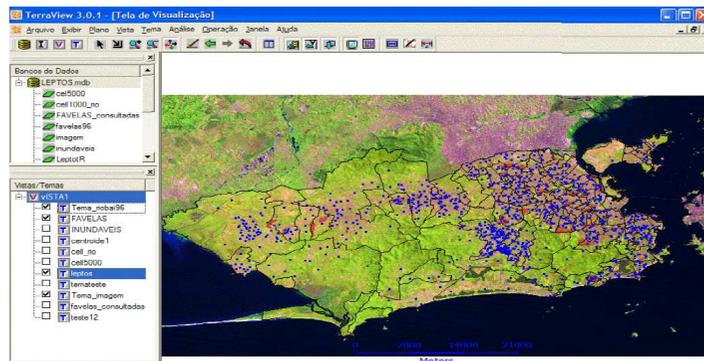
Cartografia



Uso da Terra



Monitoramento ambiental



Saúde

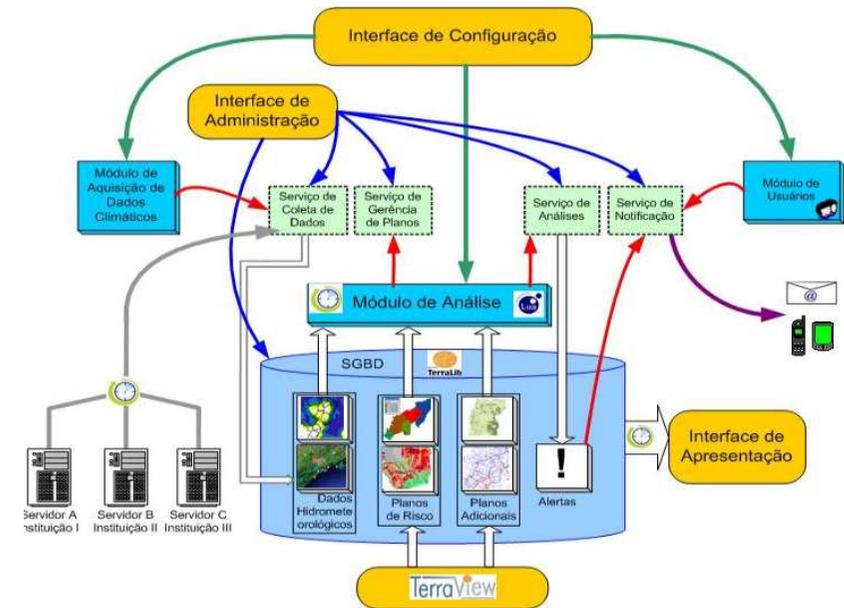


# Aplicações dos SIGs

## ■ Outras aplicações incluem:

- Cadastro urbano
- Desenvolvimento imobiliário
- Desenvolvimento econômico
- Marketing
- Navegação
- Climatologia
- Gerenciamento de aparelhos públicos
- Redes de transporte
- Biodiversidade
- Monitoramento de desastres ambientais
- ...

- SIGs tratam fundamentalmente da solução de problemas. São usados em vários níveis, desde indústrias especializadas até governos e academia





# Dados espaciais

- De acordo com algumas fontes, 80% dos dados existentes possuem uma componente espacial (pense em endereços por exemplo)
- Enormes quantidades de dados espaciais são gerados por sistemas de observação da terra: ex: dados do CBERS, LANDSAT, SRTM, etc.
- Vários serviços como MapQuest, Yahoo! Maps, Google Maps, Google
- Os serviços geográficos estão crescendo:
  - Planejamento de rotas
  - Observação e Mitigação de desastres naturais
  - Observação de crimes
  - Simulação de efeitos ambientais
  - Serviços Baseados em Localização em geral
- Em geral, é necessário integrar informações espaciais em diferentes contextos.



# Manipulação de dados geográficos

- Assumindo que existam 2 camadas representadas por:
  - Pais(nome, capital, população, geo:região) (fronteiras)
  - Linguas(lingua, geo:região) (mapa colorido)
- Operações simples, muito executadas em aplicações SIG:
  - *Projeção*: camada x {A1, ..., An} → camada  
ex., projeção sobre o atributo população
  - *Seleção*: camada x p(Ai) → camada  
ex., países com população > 30M
  - *União*: camada x camada → camada  
temas devem ter o mesmo esquema
  - *Interseção (overlay)*: camada x camada → camada  
cria uma nova camada; similar a junção espacial,  
onde objetos são ligados se suas geometrias se interceptam





# Manipulação de dados espaciais

- Seleção espacial:
  - *Consulta por janela*: obtém-se os objetos que interceptam um data área (usualmente retangular) one obtains another theme that only includes those
  - *Consulta por apontamento*: recupera os objetos cuja geometria contém um certo ponto
  - *Clipping*: extrai a porção da camada que intercepta uma certa geometria
- Outras operações
  - *Métricas*: por exemplo, distancia entre dois pontos ou regiões
  - *Topológicas*: ex. objetos que são vizinhos
- Outras operações típicas incluem interpolação, localização, alocação, etc.
- Operações podem ser formalmente descritas usando uma álgebra sobre mapas



# SIG

- Uma **camada** (ou **mapa**, ou **tema**) de informação geospacial corresponde a um tópico em particular
- Quando uma camada ou um mapa é desenhado, o usuário tipicamente vê um mapa, com cores, em uma certa escala, com uma certa legenda, etc.
- A construção de um modelo conceitual para as camadas de interesse é tipicamente o primeiro passo de uma aplicação em um SIG.  
  
    camada = {objeto geográfico}  
  
    objeto geográfico = (descrição, geometria) // objeto atômico  
                          | (descrição, {objeto geográfico}) // objeto complexo
- Um camada é um conjunto de objetos geográficos homogêneos, ou seja, que possuem a mesma estrutura, representada de acordo com o modelo lógico do SIG sendo utilizado.
- A representação da geometria e as relações geométrica requerem algum modelo espacial que inclua os tipos básicos: pontos (0-D), linhas (1-D) e regiões (2-D)



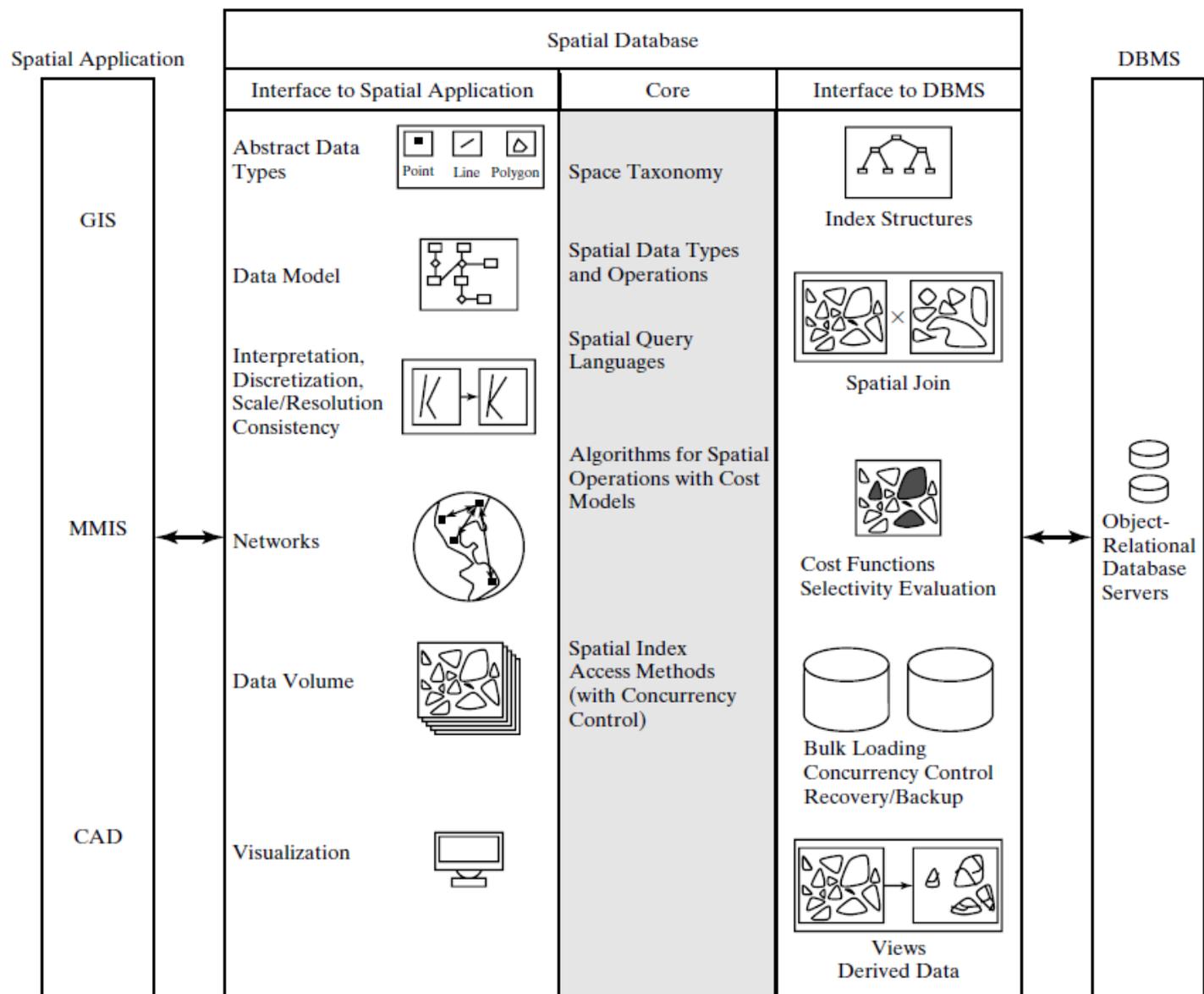
# Relação SIG e Banco de Dados

- Um SIG necessita manipular tanto a componente alfa numérica quanto a espacial.
- Inicialmente os SIGs eram construídos sobre formatos proprietários de arquivos para a componente espacial.
- Também é possível armazenar a componente espacial usando apenas os tipos convencionais do banco de dados:
  - Por exemplo, um polígono pode ser descrito usando uma sequência de linhas, e cada linha descrita por seu ponto inicial e final
- Modelo pode ser considerado inapropriado por razões como:
  - Princípio da independência de dados é violado. É necessário conhecer a estrutura interna de representação para formular consultas espaciais
  - Baixa performance.
  - Não é user-friendly
  - Difícil definir tipos particulares
  - Impossível expressar certos cálculos geométricos, como testes de adjacência, consultas por apontamento ou por janela



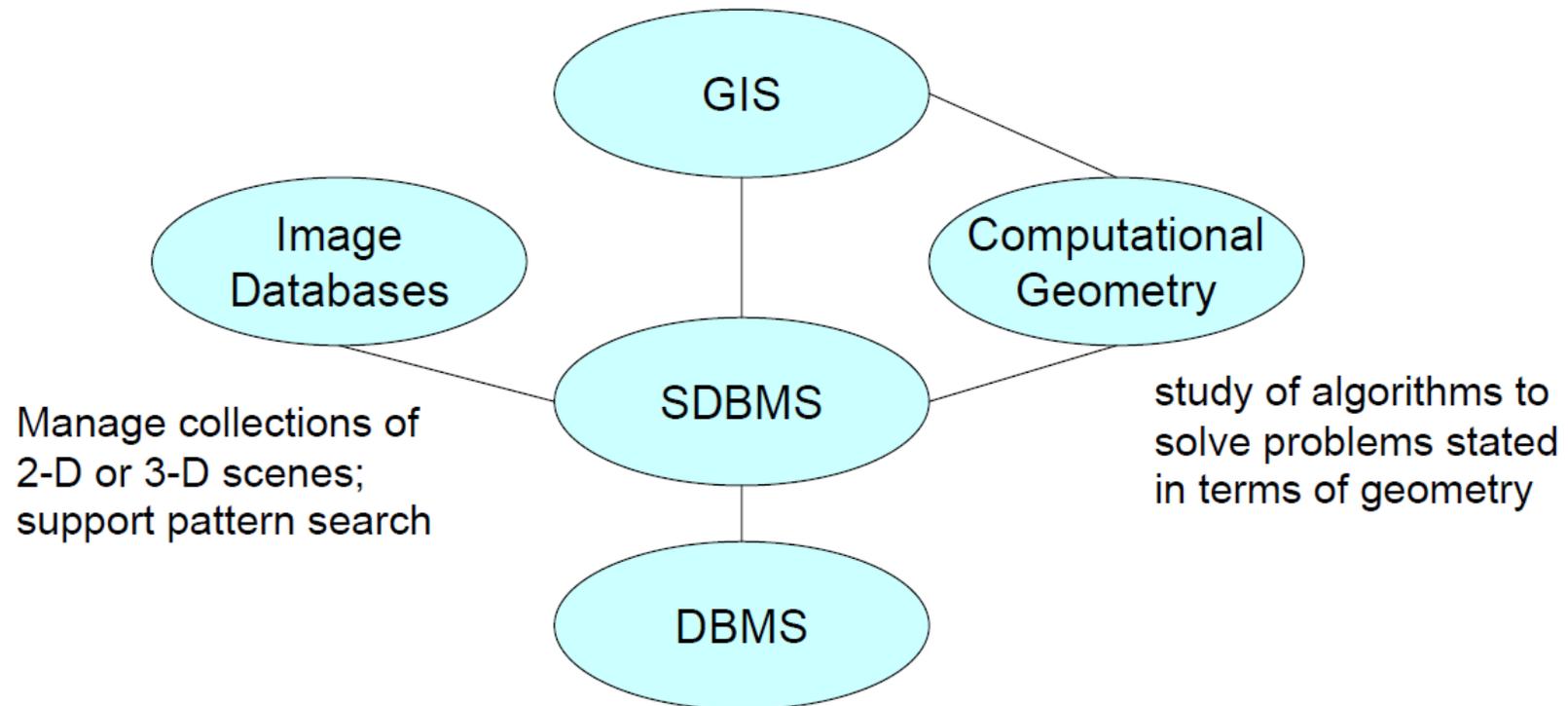
# Modelo em 3 camadas

Fonte: Shekkar e Chawla, 2003





# Contexto



Fonte: Gertz, 2008.



MINISTÉRIO DA CIÊNCIA E TECNOLOGIA  
**INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS**

# SGBDs



# Introdução

Um **Banco de Dados** (BD) é uma coleção volumosa de dados relacionados entre si, em um ambiente computacional

Esses dados são persistentes, ou seja, sobrevivem a falhas de hardware ou software

Um **Sistema Gerenciador de Banco de Dados** (SGBD) é um sistema computacional que gerencia a estrutura do banco de dados e controla o acesso aos dados armazenados



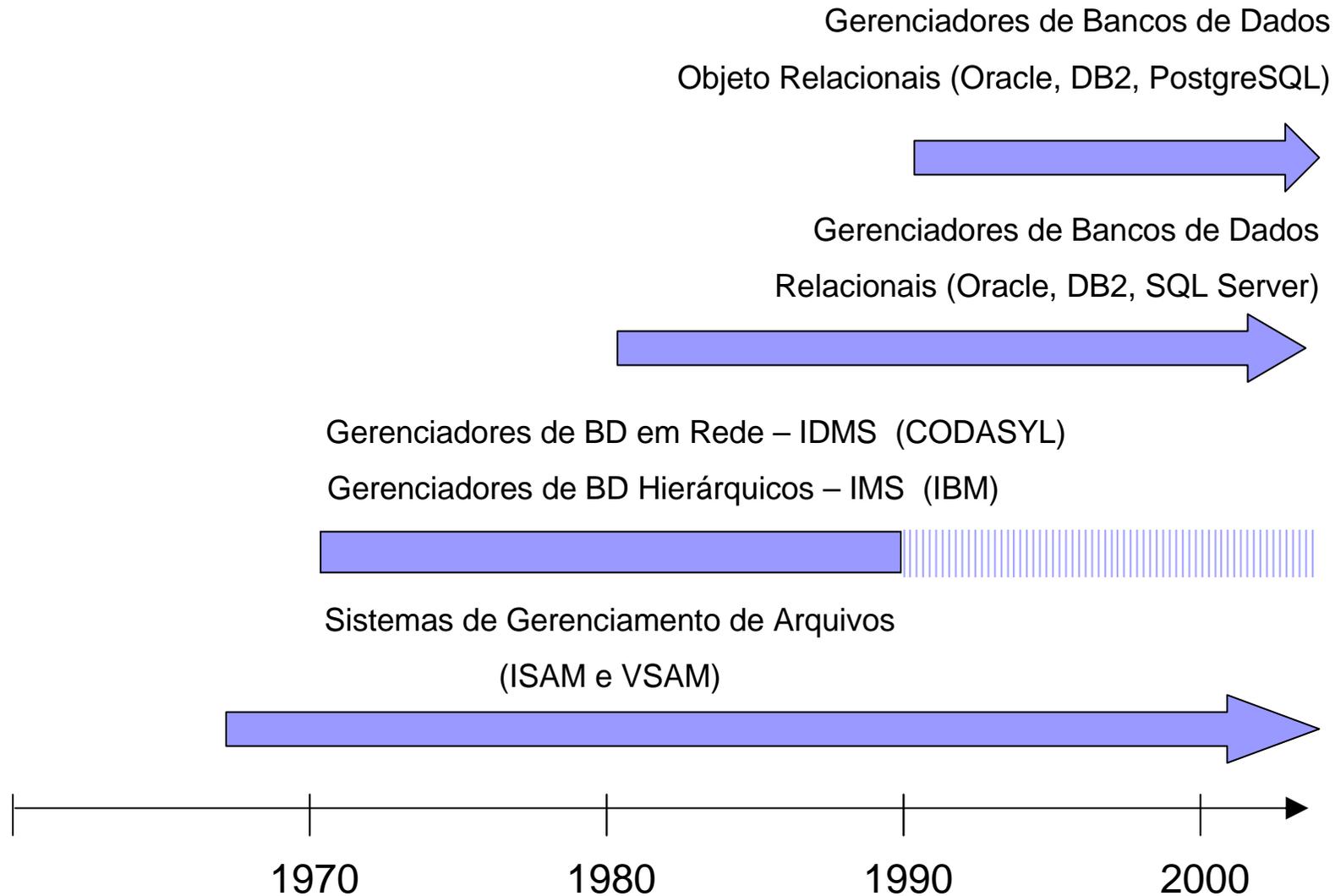


# Introdução

- Um **SGBD** facilita as seguintes tarefas:
  - **Definição** do BD: tipos de dados, estruturas e restrições que devem ser consideradas
  - **Construção** do BD: inserção propriamente dos dados no meio persistente;
  - **Manipulação** do BD
  - **Consulta** ao BD: recuperar uma informação específica
  - **Manutenção** ao BD: alterar uma informação armazenada
- São requisitos de um SGBD:
  - Facilidade de uso
  - Correção
  - Facilidade de manutenção
  - Confiabilidade
  - Segurança
  - Desempenho



# Histórico





# Sistemas de Arquivos Convencionais

- Características:
  - Vários arquivos, diferentes programas para recuperar informações
  
- Desvantagens:
  - Redundância e inconsistência de dados
  - Dificuldade de acessar informação (programa pode não estar disponível)
  - Anomalias de acesso concorrente (multiusuários)
  - Problemas de segurança e integridade
  - Isolamento de dados



# Sistemas de Banco de Dados

- Características:
  - Coleção de arquivos inter-relacionados + Conjunto de programas para armazenar, gerenciar e acessar esses arquivos.
  
- Vantagens
  - Consistência e integridade dos dados
  - Segurança
  - Controle de acesso concorrente (multi-usuário)
  - Backup e recuperação de falhas



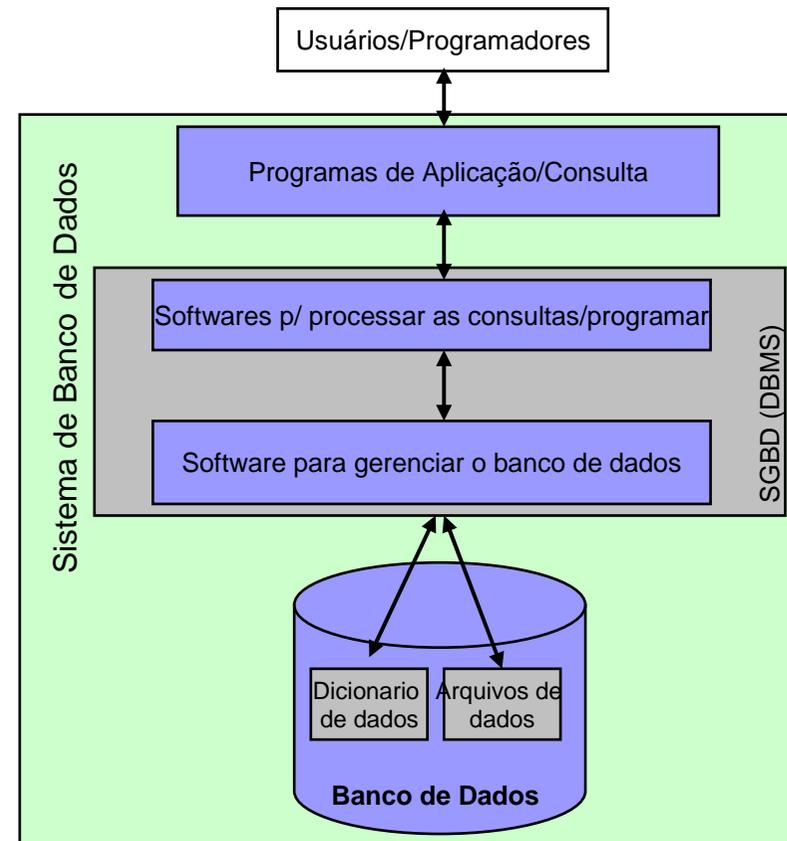
# Características de um SGBD

- Um **SGBD** facilita as seguintes tarefas:
  - **Definição** do BD: tipos de dados, estruturas e restrições que devem ser consideradas
  - **Construção** do BD: inserção propriamente dos dados no meio persistente;
  - **Manipulação** do BD
  - **Consulta** ao BD: recuperar uma informação específica
  - **Manutenção** ao BD: alterar uma informação armazenada
- São requisitos de um SGBD:
  - Facilidade de uso
  - Correção
  - Facilidade de manutenção
  - Confiabilidade
  - Segurança
  - Desempenho



# Um SGBD oferece

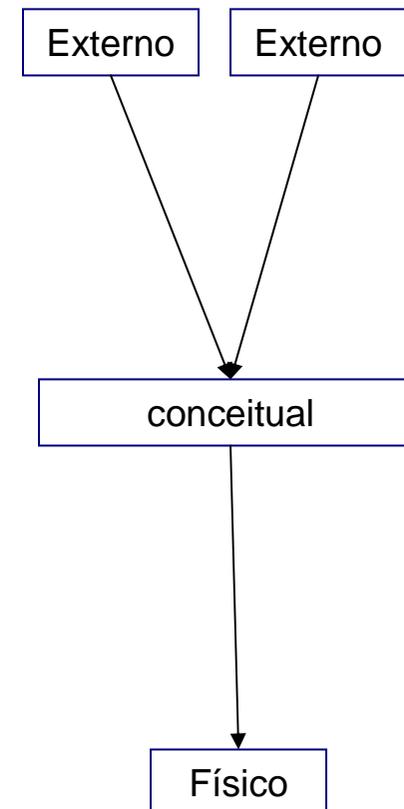
- Interfaces baseadas em modelo de dados de alto nível tanto para a definição da estrutura da base quanto para sua consulta
- Mecanismos que garantem restrições de integridade (ex. triggers, assertions)
- Atomicidade-consistência-integridade-durabilidade) (ex. controle de concorrência, subsistema de recuperação)
- Controle de acesso
- Métodos de acesso e armazenamento eficientes (ex. otimização de consultas)





# Especificação em níveis

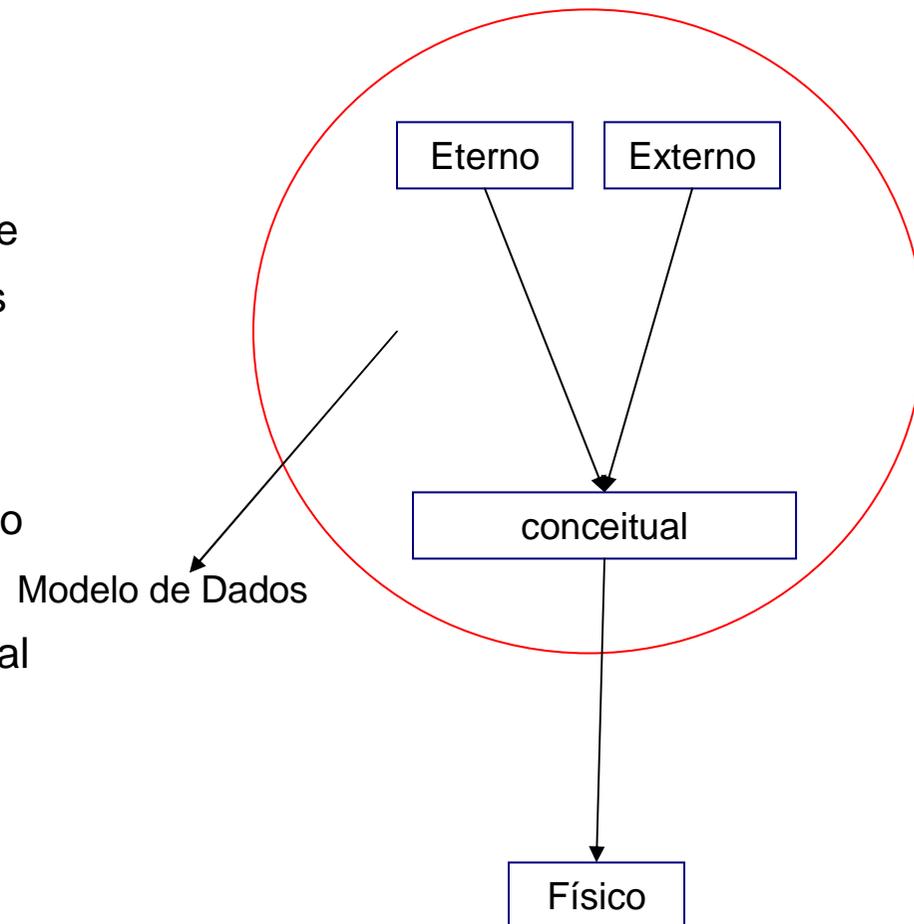
- nível externo:
  - especificação da organização conceitual do BD, vista por um grupo de usuários
- nível conceitual:
  - especificação da organização conceitual do BD, ou seja, o que o BD armazena
- nível físico ou interno:
  - especificação das estruturas de armazenamento do BD, ou seja, como o BD está armazenado





# Vantagens da especificação em níveis

- Facilidade de manutenção
- Independência física (dos dados)
  - permite modificar as estruturas de armazenamento sem impactar as aplicações
- Independência lógica (dos dados)
  - separação entre esquema externo e esquema conceitual permite modificar a organização conceitual com impacto mínimo nas aplicações (construídas sobre os esquemas externos)





# Modelos de dados

- Conjunto de conceitos usados para representar os dados, os relacionamentos entre esses dados e as restrições de consistência
- No processo de modelagem é necessário construir uma abstração dos objetos e fenômenos do mundo real.

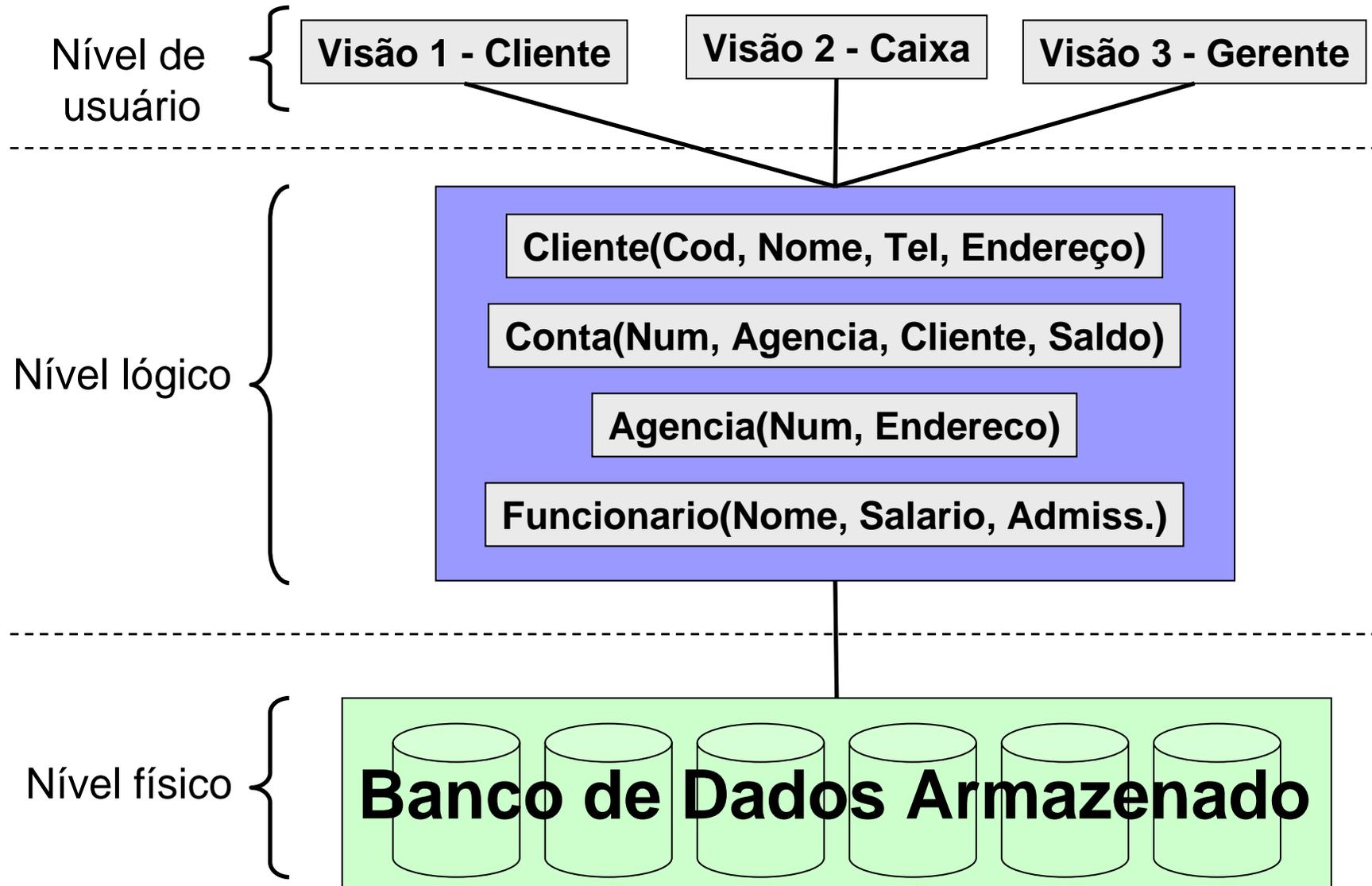


# Abstração de dados

- Nível de visões (de usuário)
  - Alto nível de abstração
  - Diferentes usuários podem ter diferentes visões do BD
  
- Nível lógico ou conceitual
  - Descreve *quais* dados estão armazenados e as relações entre eles
  
- Nível físico
  - Descreve *como* os dados estão armazenados
  - Baixo nível de abstração
  - Estruturas complexas e detalhadas



# Abstração de dados





# Projeto de um Banco de Dados

- Começa com um levantamento de requisitos
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
- Termina com o projeto físico do banco



# Projeto de um Banco de Dados

- Começa com um levantamento de requisitos
- Modelagem conceitual
- Modelagem lógica
- Refinamento e projeto físico
- Termina com o projeto físico do banco



# Modelo Entidade-Relacionamento (E-R)

- Introduzido em 1976 por Peter Chen, é a abordagem mais adotada para modelagem conceitual de dado. Objetivo é facilitar o projeto do banco de dados representando sua estrutura lógica
- Definição: modelo baseado na percepção do mundo real como um conjunto de objetos chamados entidades e pelo conjunto de relacionamentos entre esses objetos.
- Ao longo do tempo diferentes versões de foram sendo propostas para representar diferentes conceitos.



# Modelo Entidade-Relacionamento (E-R)

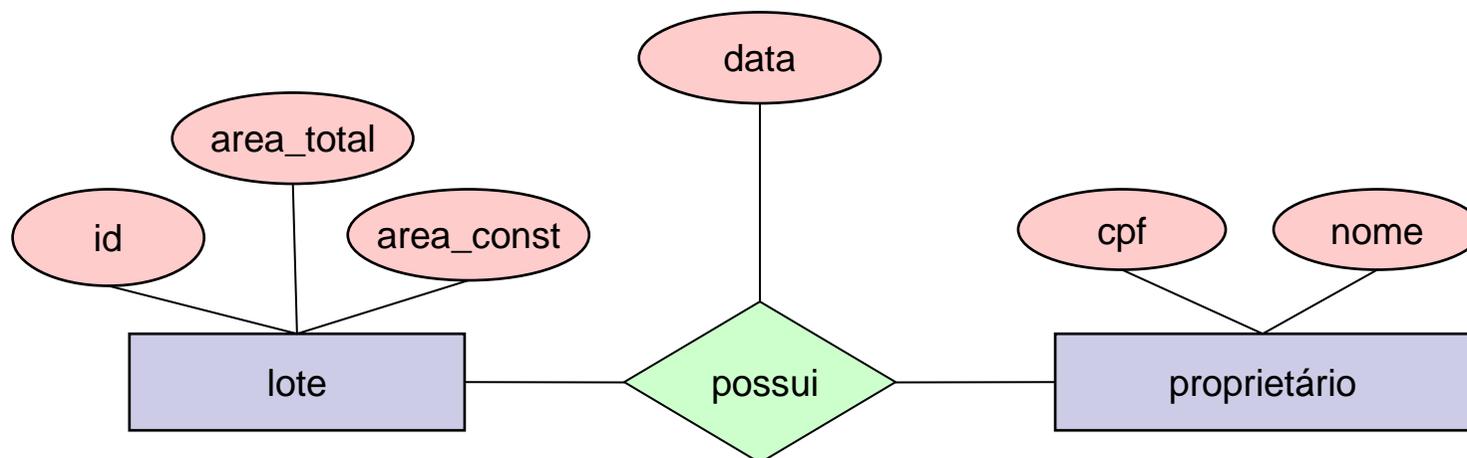
- Conceitos básicos:
  - Entidades
    - Objetos básicos do mundo real
    - Um conjunto de entidades agrupa entidades do mesmo tipo
  - Relacionamentos
    - Associação entre conjuntos de entidades
  - Atributos
    - Associados a entidades e a relacionamentos
    - Uma entidade é representada por um conjunto de atributos
    - Cada atributo possui um domínio



# Modelo Entidade-Relacionamento (E-R)

## ■ Representação:

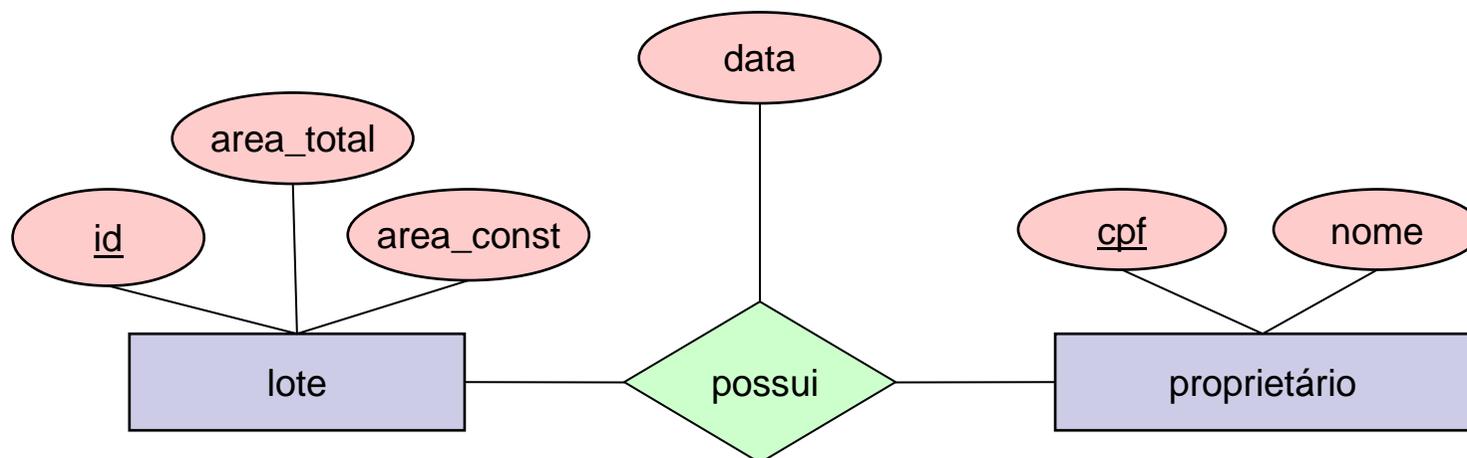
- Retângulos: conjunto de entidades
- Elipses: atributos
- Losangos: relacionamentos





# Modelo Entidade-Relacionamento (E-R)

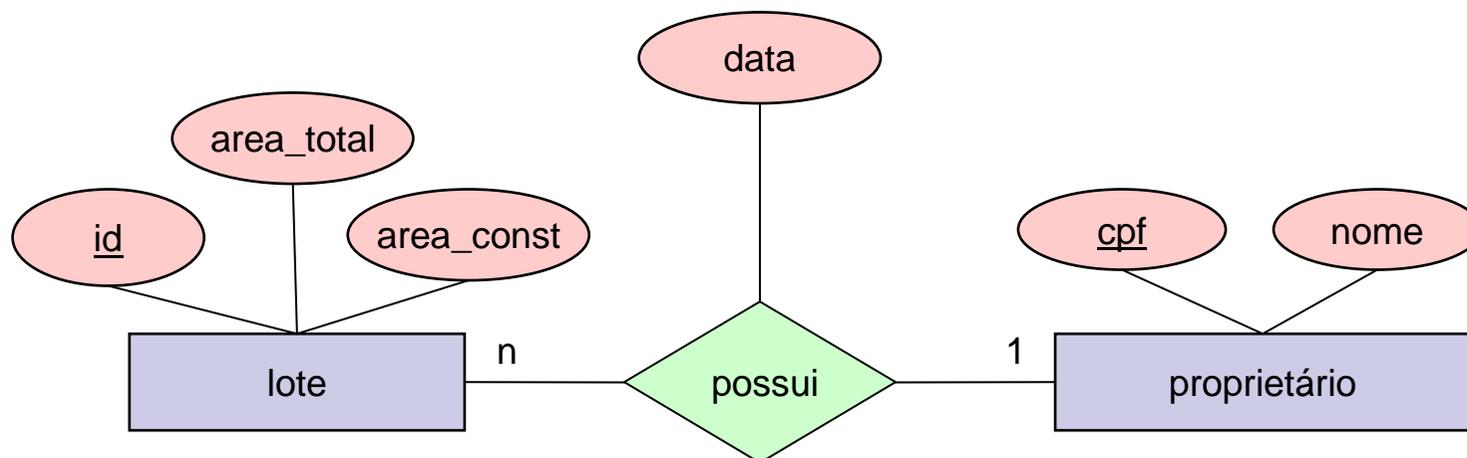
- Atributo identificador
  - Atributo que vai identificar unicamente cada instância da entidade (chave primária)
  - Devem ser sublinhados





# Modelo Entidade-Relacionamento (E-R)

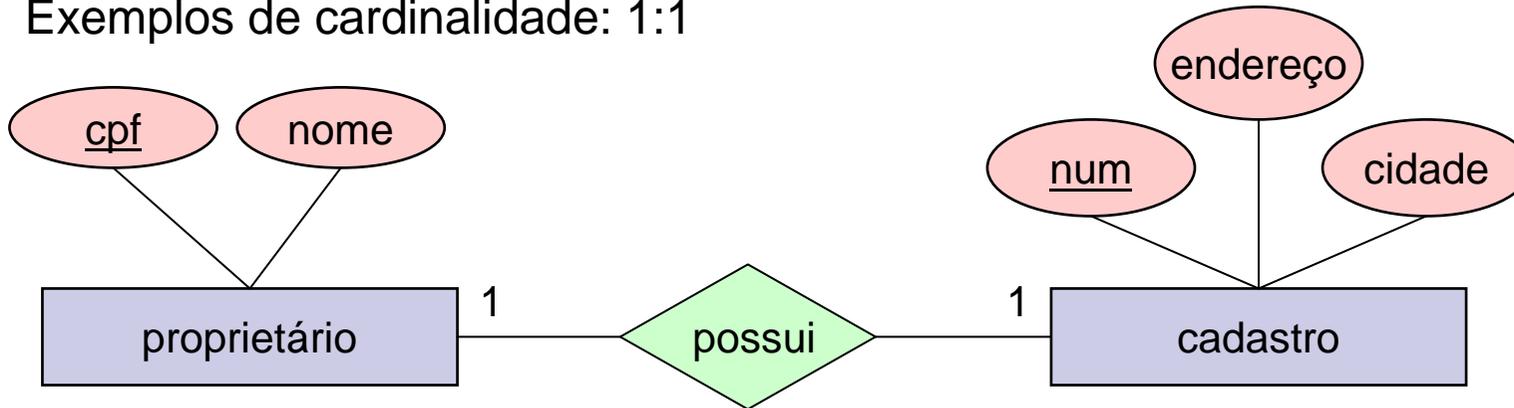
- Cardinalidade:
  - expressa o número de entidades que uma entidade pode estar associada com
  - 1:1, 1:n, n:1, n:n





# Modelo Entidade-Relacionamento (E-R)

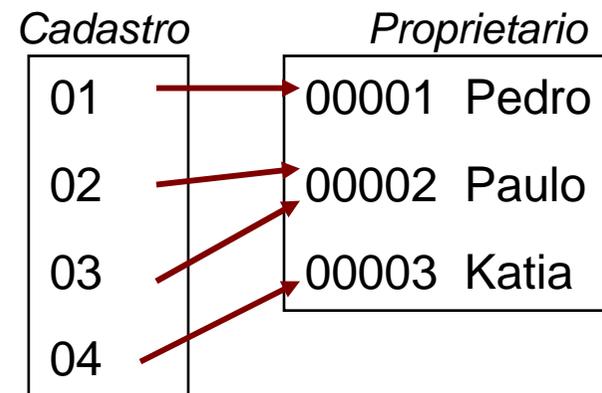
- Exemplos de cardinalidade: 1:1



Relação correta



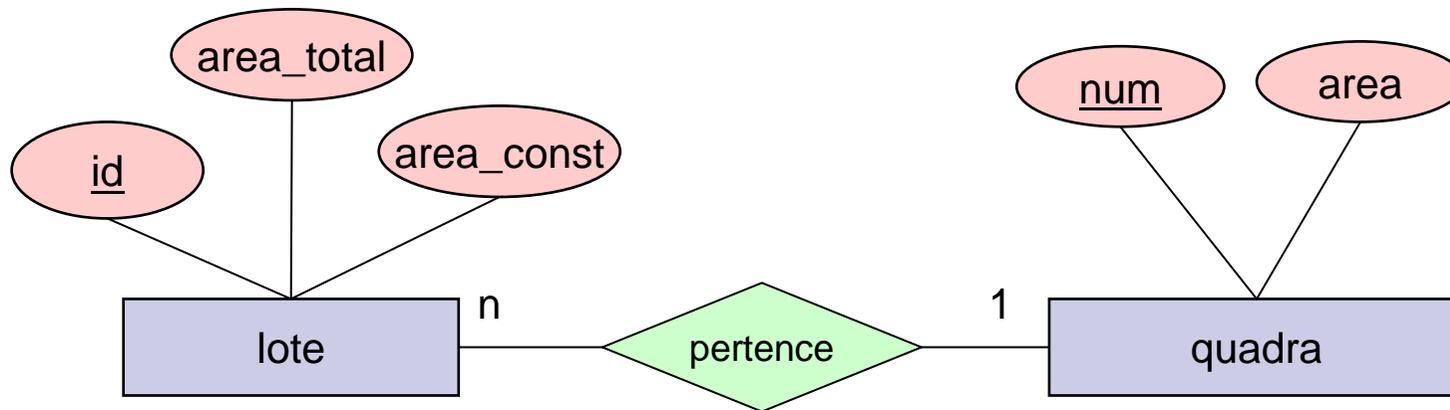
Relação Incorreta



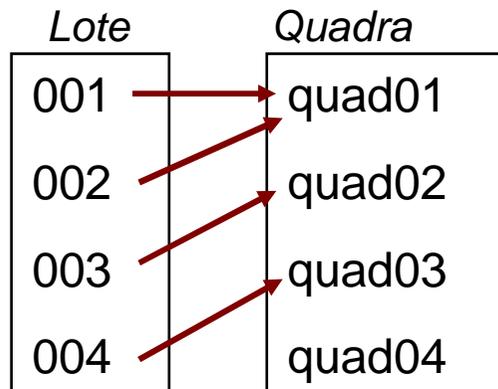


# Modelo Entidade-Relacionamento (E-R)

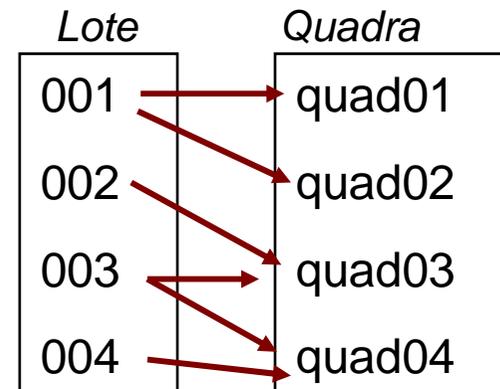
- Exemplos de cardinalidade: 1:N



Relação correta



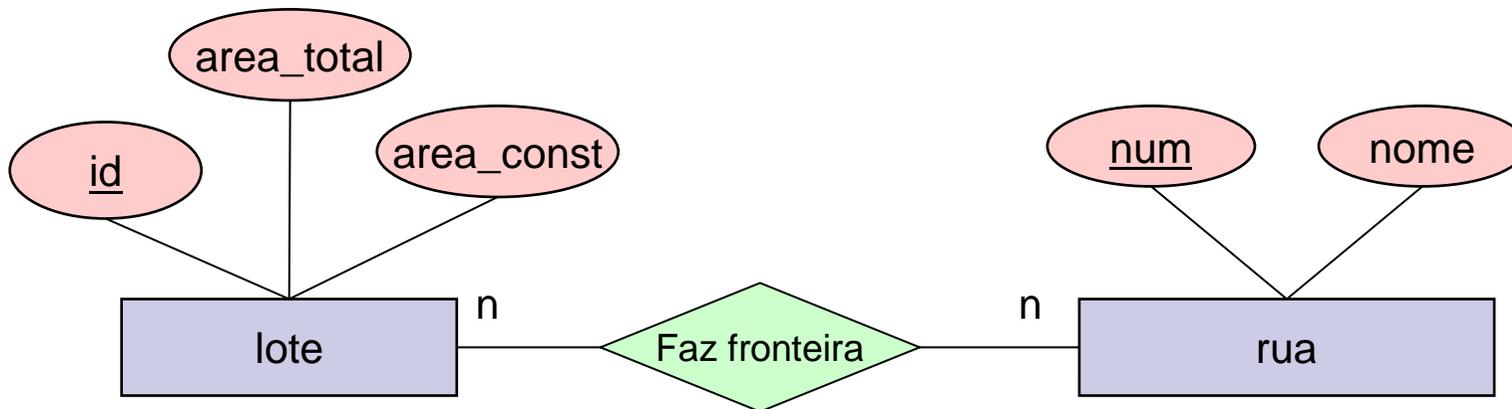
Relação incorreta



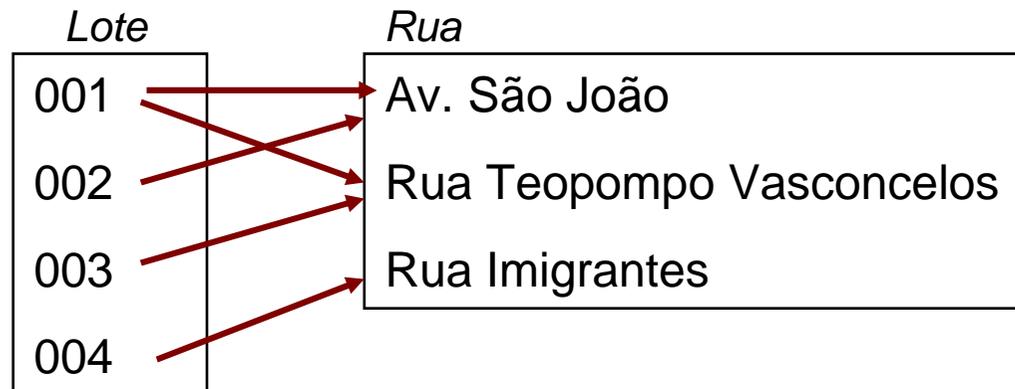


# Modelo Entidade-Relacionamento (E-R)

- Exemplos de cardinalidade: N:N



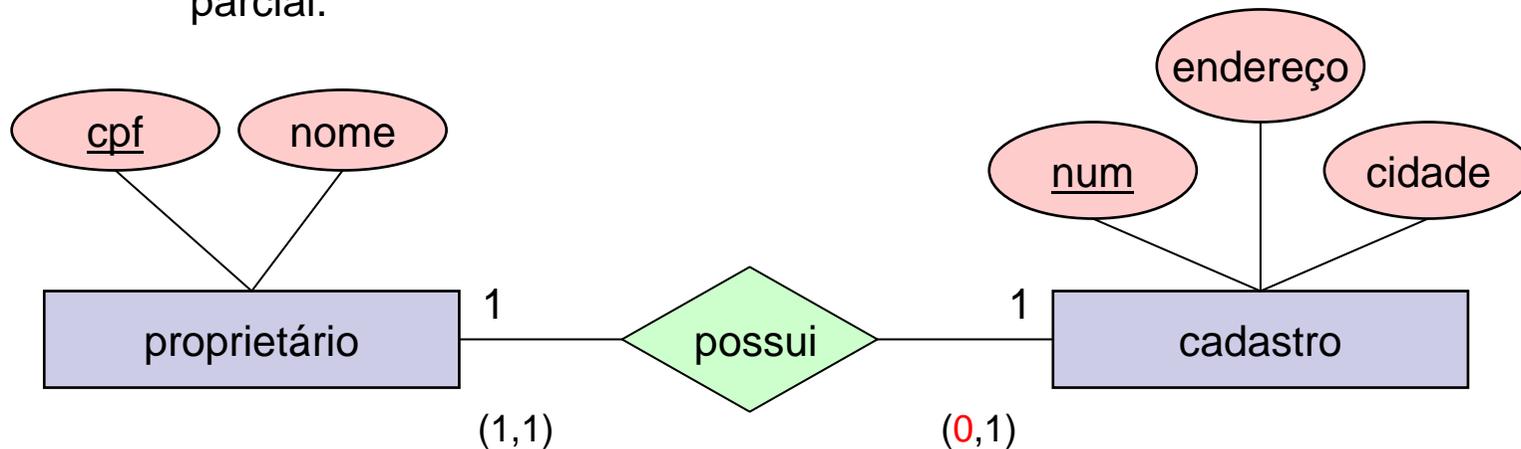
Relação correta





# Modelo Entidade-Relacionamento (E-R)

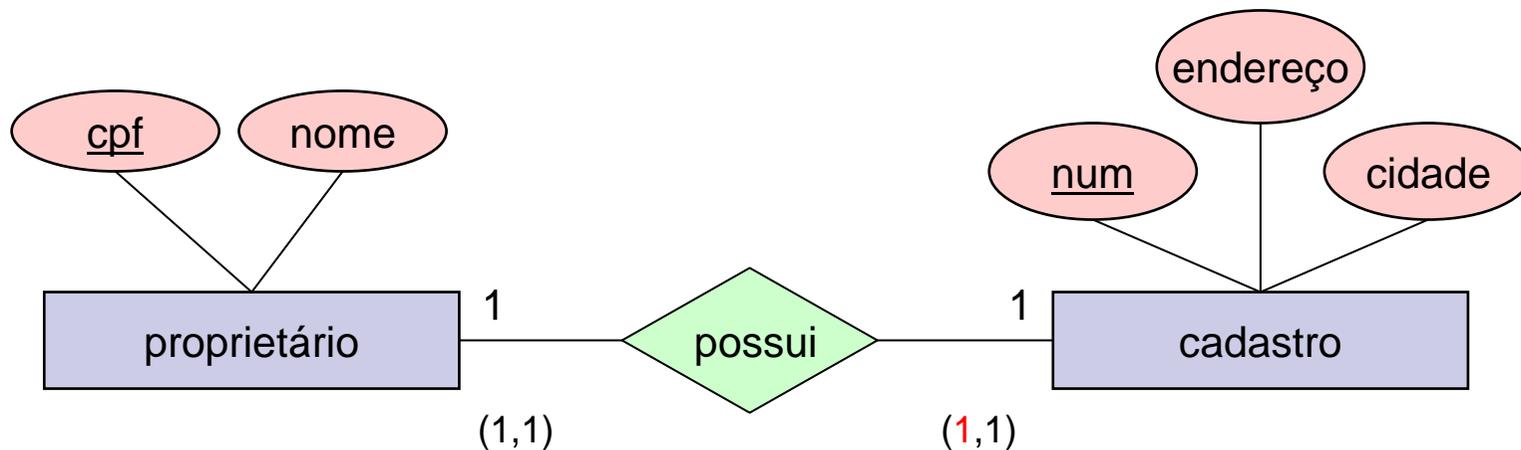
- Relacionamento parcial:
  - Quando não existe obrigatoriedade de todas as entidades de um conjunto participarem do relacionamento
    - Ex. Podem existir proprietários sem cadastro
    - A participação da entidade “proprietário” no relacionamento “possui” é parcial.





# Modelo Entidade-Relacionamento (E-R)

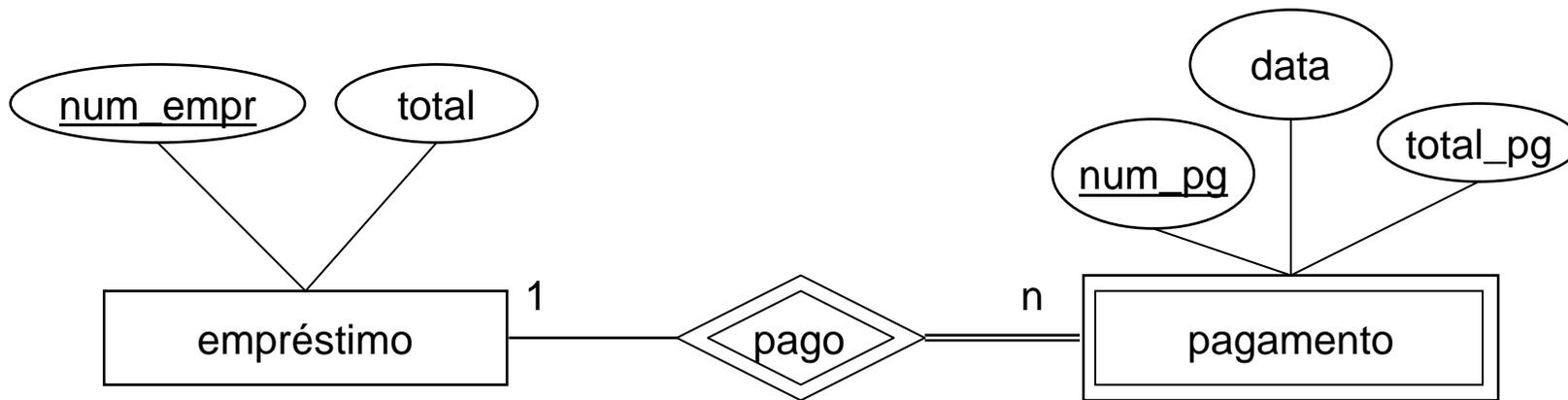
- Relacionamento total (dependência de existência):
  - Quando todas as entidades de um conjunto participam obrigatoriamente do relacionamento
    - Ex. Todo cadastro tem que estar associado a um proprietário.
    - A participação da entidade “cadastro” no relacionamento “possui” é total.





# Modelo Entidade-Relacionamento (E-R)

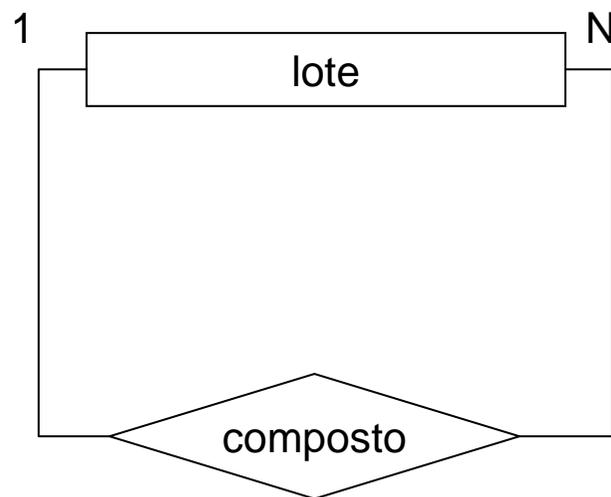
- Entidade fraca:
  - Quando a existência dessa entidade depende da existência de uma outra entidade (dependência de existência).





# Modelo Entidade-Relacionamento (E-R)

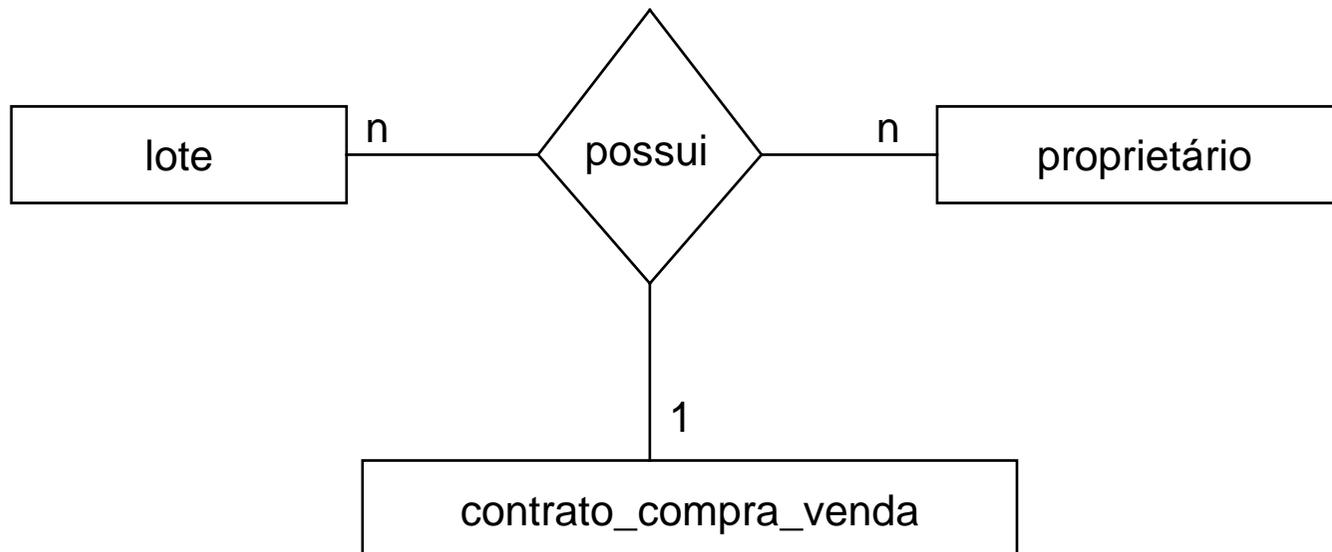
- Auto relacionamento:
  - Associa entidades de um conjunto a entidades desse mesmo conjunto
    - Ex. Um lote pode ser composto por outros lotes





# Modelo Entidade-Relacionamento (E-R)

- Relacionamento múltiplo:
  - Associa mais de dois conjuntos de entidades
    - Ex. Um proprietário pode possuir vários lotes, cada um com um contrato de compra e venda diferente.





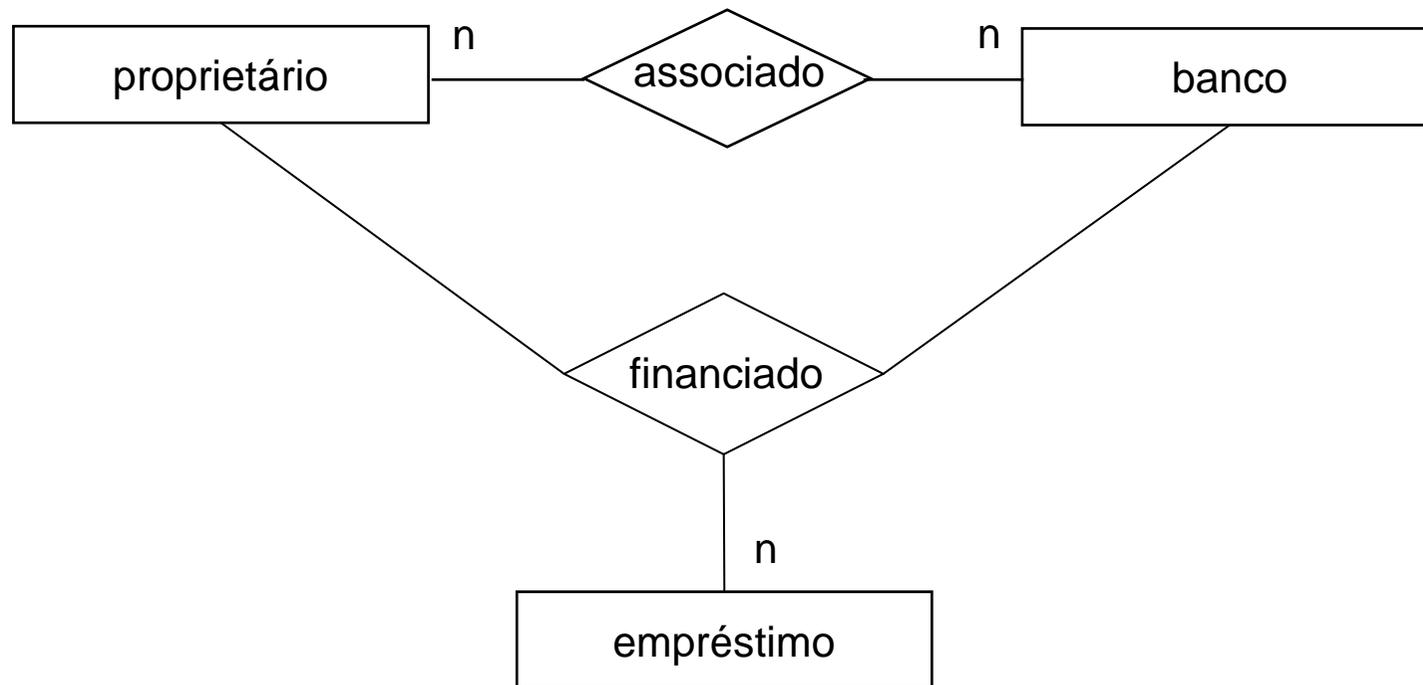
# Modelo Entidade-Relacionamento (E-R)

- Agregação:
  - Uma limitação do ER: expressar relacionamentos entre relacionamentos
  - Agregação é uma abstração na qual os relacionamentos são tratados como conjunto de entidades de nível superior, permitindo o relacionamento entre relacionamentos.



# Modelo Entidade-Relacionamento (E-R)

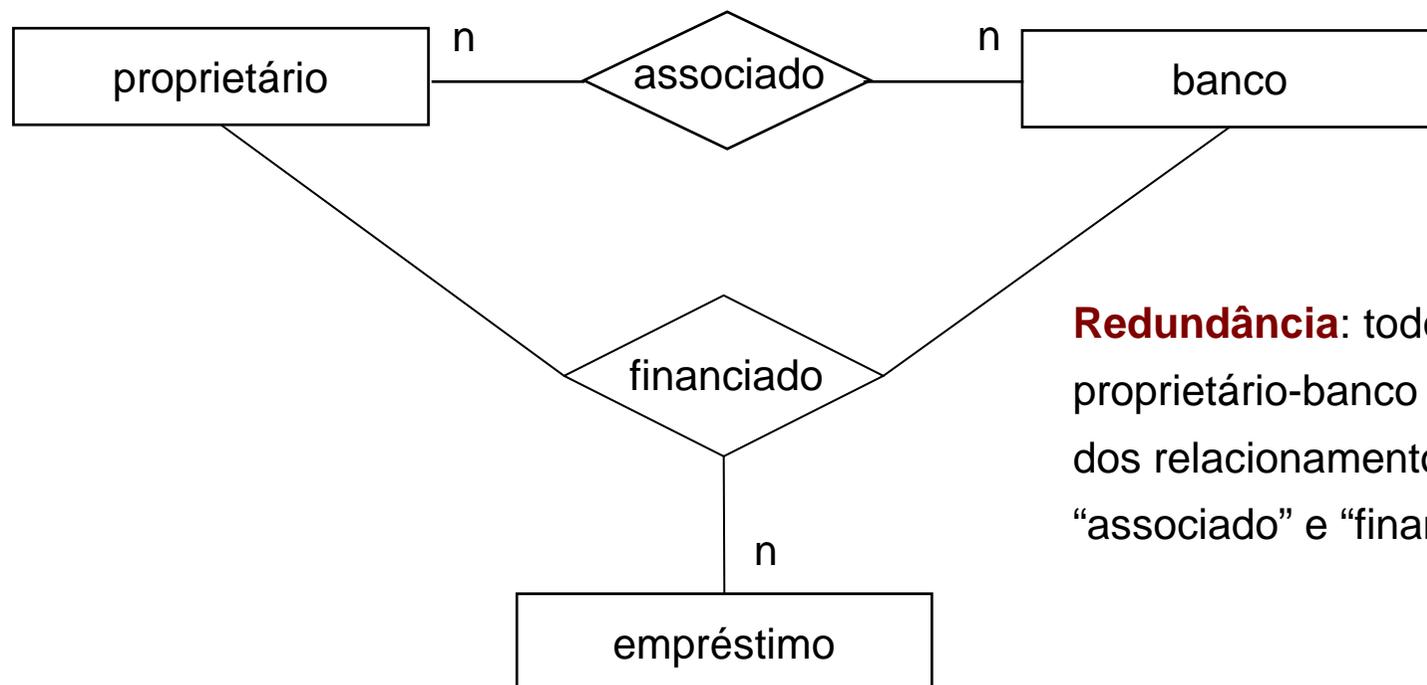
- Ex.: um proprietário pode estar associado a vários bancos. Pra cada banco associado, o proprietário pode ter um ou mais empréstimos.





# Modelo Entidade-Relacionamento (E-R)

- Ex.: um proprietário pode estar associado a vários bancos. Pra cada banco associado, o proprietário pode ter um ou mais empréstimos.

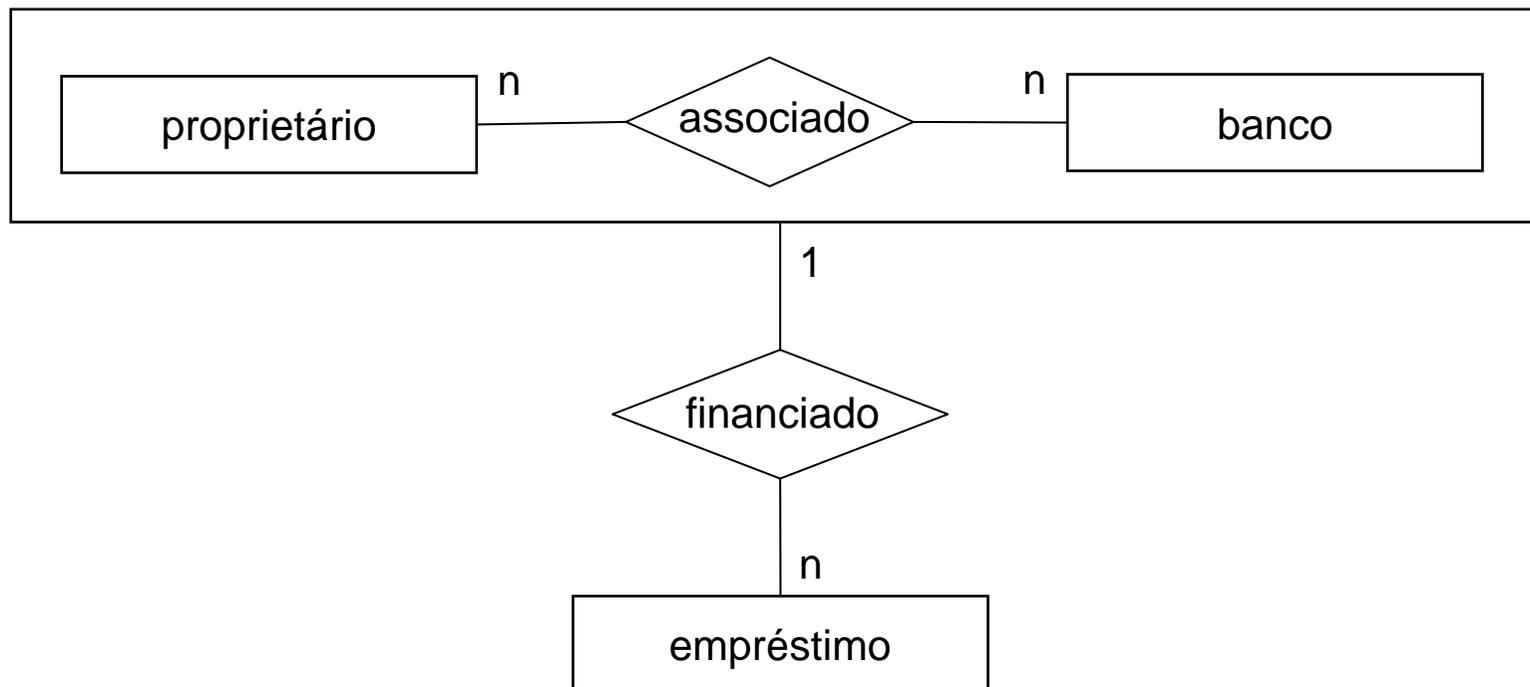


**Redundância:** todo par proprietário-banco faz parte dos relacionamentos “associado” e “financiado”



# Modelo Entidade-Relacionamento (E-R)

- Agregação:
  - Ex.: um proprietário pode estar associado a vários bancos. Pra cada banco associado, o proprietário pode ter um ou mais empréstimos.





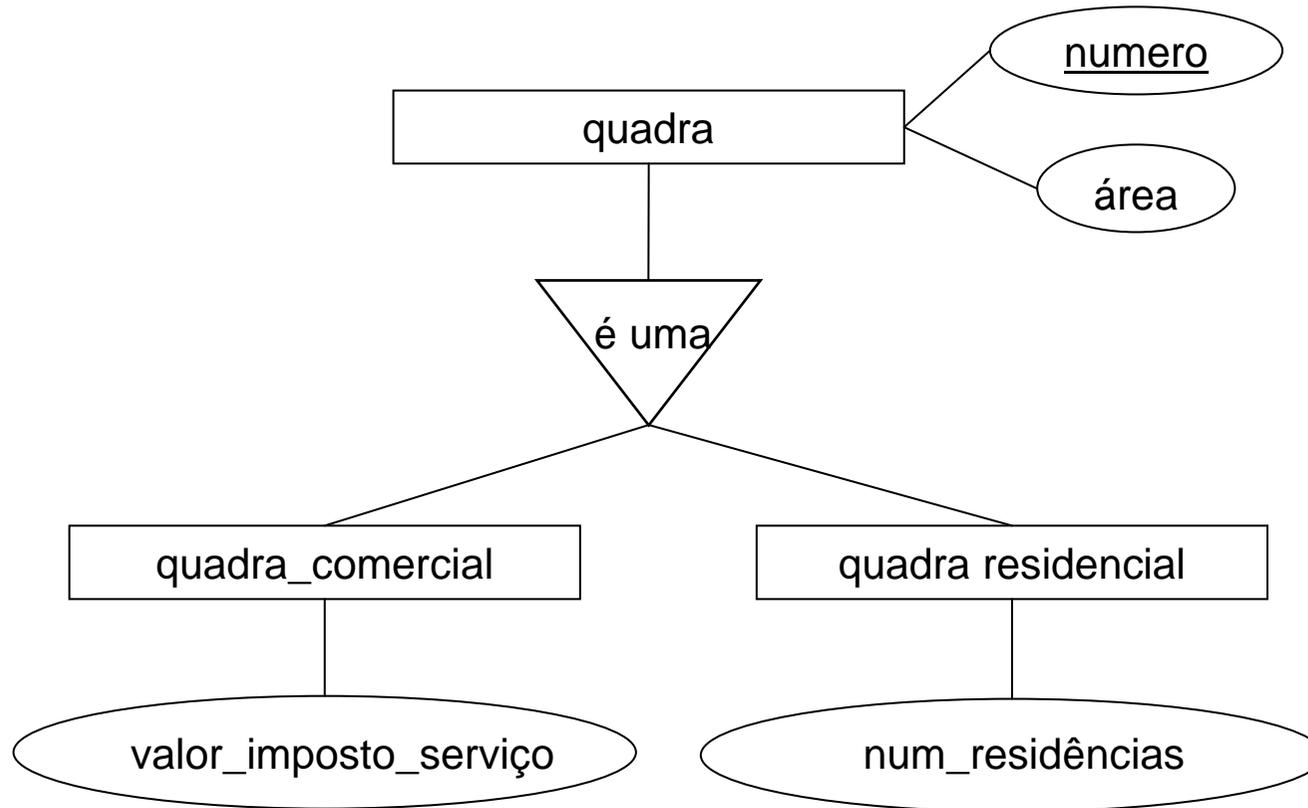
# Modelo Entidade-Relacionamento (E-R)

- Generalização
  - Enfatiza a semelhança entre diferentes tipos de entidades e abstrai suas diferenças.
  
- Especialização:
  - Subdivide entidades semelhantes em conjuntos de entidades mais específicas e enfatiza suas diferenças.



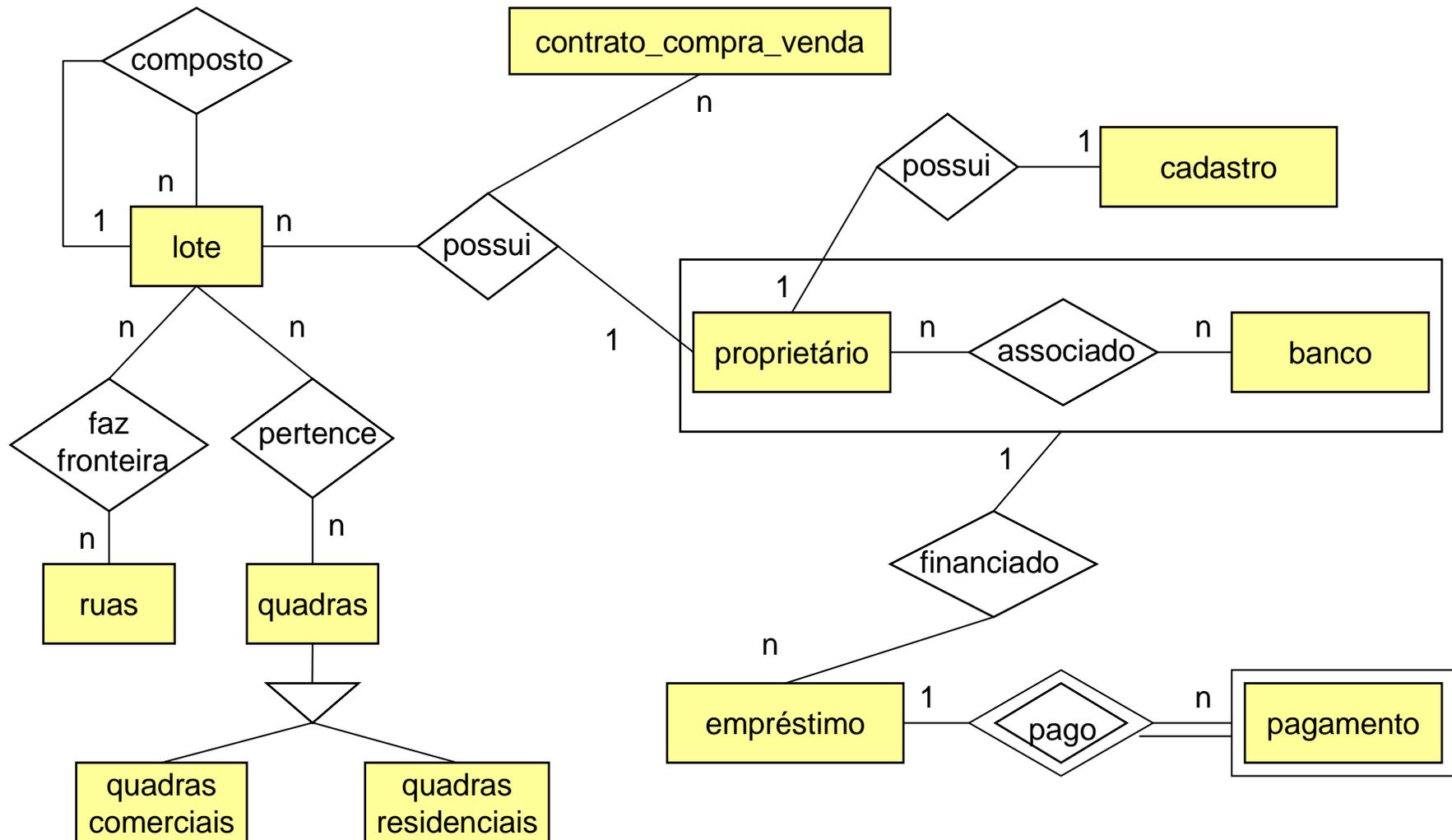
# Modelo Entidade-Relacionamento (E-R)

- Generalização e Especialização
  - Ex.: Uma quadra pode ser do tipo comercial ou residencial



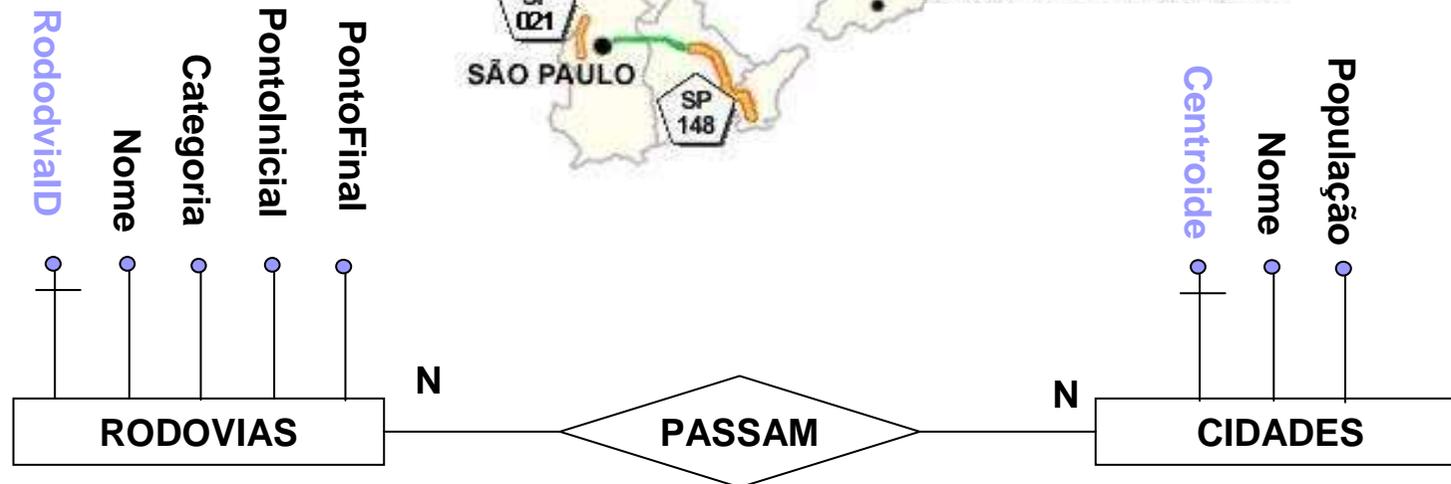


# Modelo Entidade-Relacionamento (E-R)





# Modelo Entidade-Relacionamento (E-R)





# O que é um bom modelo?

- Facilita a descrição de consultas corretas e compreensíveis
- Os elementos do modelo tem um significado intuitivo
- O modelo é o mais simples possível, mas não mais simples que isso
- É suficientemente abstrato para não sofrer alterações face a alterações menores no domínio do problema
- Se o domínio do problema se altera significativamente é fácil modificar o modelo de dados (suficientemente flexível)
- Mais tarde é necessário considerar o impacto do modelo de dados na eficiência das operações no banco de dados



# Decisões de projeto

- Um conceito deve ser modelado como entidade ou atributo?
- Um conceito deve ser modelado como uma entidade ou um relacionamento?
- Identificação dos relacionamentos: binário, ternário? Agregação?
- Restrições no modelo ER:
- Muita semântica pode (e deve) ser capturada mas algumas restrições não podem ser capturadas nos diagramas ER



## Exemplo: Entidade X Atributo

- Exemplo: o *endereço* de um *empregado* deve ser modelado como um atributo ou uma entidade?
- Depende do problema:
  - se um empregado pode ter mais que um endereço então ele deve ser uma entidade separada porque atributos só possuem um valor
  - se faremos consultas sobre partes do endereço (ex. cidade) então ele deve ser uma entidade separada porque atributos devem ser atômicos



# Modelo Relacional

- O modelo relacional foi inventado por E.F. (Ted) Codd como um modelo geral de dados (~1970)
- Baseia-se no princípio de que todos os dados são representados matematicamente como relações *n-árias*, onde uma *relação n-ária* é um subconjunto do produto cartesiano de *n* domínios
- Dados são manipulados através de um *cálculo* ou *álgebra relacional*
- Primeiro modelo de dados para Sistemas de Bancos de Dados comerciais
- É usado na maioria das aplicações para dados tradicionais
- Os SGBDs relacionais são baseados no modelo relacional
  - IBM, Informix, Microsoft, Oracle, Sybase, etc.



# Modelo Relacional

- Banco de Dados Relacional
  - Coleção de tabelas, compostas por colunas e linhas inter-relacionadas

## PROPRIETARIO

CPF	NOME	RUA	NUMERO	BAIRRO
08940256	JOÃO DA SILVA	SAO JOAO	180	CENTRO
03727298	HENRIQUE CARDOSO	IMIGRANTE	1700	VILA 12
97260089	JOSÉ DE SOUZA	SAO JOAO	35	CENTRO

## LOTE



NUMERO	PROPRIETARIO_CPF	AREA_TOTAL	AREA_CONST
00001	08940256	400.000	0
00003	03727298	150.000	75.00
00039	03727298	500.000	0



# Modelo Relacional

- Características:
  - Consiste em um conjunto de tabelas ou *relações* formadas por linhas e colunas
  
  - Cada coluna (campo):
    - Representa um atributo
    - Está associada a um domínio (conjunto de valores permitidos)
  
  - Cada linha (registro ou *tupla*):
    - Representa um relacionamento entre um conjunto de valores para cada atributo



# Modelo Relacional

- Conceito de *relação*
  - Define uma tabela do banco de dados
  - Domínio de um atributo: conjunto de possíveis valores

- Ex.:

$$D_1 = \{ x \in \mathfrak{R} \mid x \geq -5 \text{ e } x \leq 5 \}$$

$$D_2 = \{ y \in \mathfrak{R} \mid y \geq 0 \}$$



# Modelo Relacional

- Conceito de *relação*

- Dados os domínios  $D_1, D_2, \dots, D_n$  não necessariamente distintos, uma relação é definida como:

$$R = \{ (d_1, d_2, \dots, d_n) \mid d_1 \in D_1, d_2 \in D_2, \dots, d_n \in D_n \}$$

- O conjunto  $(d_1, d_2, \dots, d_n)$  de valores ordenados define uma *tupla*
- Uma relação é o conjunto de *n-tuplas* ordenadas, onde  $n$  define o grau da relação



# Modelo Relacional

- Exemplo de *relação*

## PROPRIETARIO

<b>CPF</b>	<b>NOME</b>	<b>RUA</b>	<b>NUMERO</b>	<b>BAIRRO</b>
08940256	JOÃO DA SILVA	SAO JOAO	180	CENTRO
03727298	HENRIQUE CARDOSO	IMIGRANTE	1700	VILA 12
97260089	JOSÉ DE SOUZA	SAO JOAO	35	CENTRO

### Atributo

cpf

nome

rua

numero

bairro

### Domínio

inteiro longo positivo

conjunto de caracteres

conjunto de caracteres

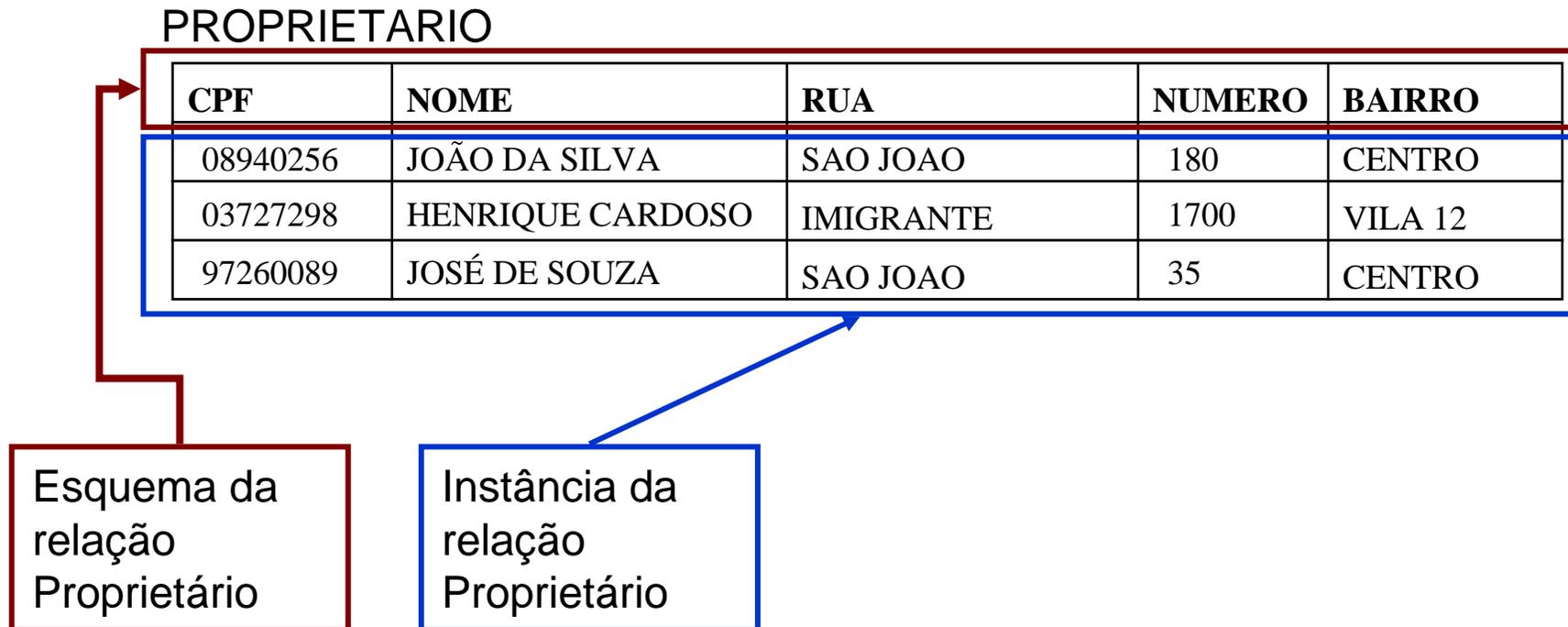
inteiro positivo

conjunto de caracteres



# Modelo Relacional

- Esquema x instância de *relação*





# Modelo Relacional

- Restrições de integridade
  - Uma das funcionalidades básicas que todo SGBD deve oferecer
  - É uma regra de consistência de dados que é garantida pelo SGBD
  - Alguns tipos de Restrições:
    - **Restrição de domínio**
    - **Restrição de chave**
    - **Integridade Referencial**



# Modelo Relacional

- Super-chave:
  - Conjunto de um ou mais atributos que permitem identificar cada registro da tabela como único.
  
- Chave candidata:
  - corresponde a super-chave mínima, ou seja, não existe sub-conjunto da chave candidata.
    - { cpf, nome } chave candidata?
    - { cpf } chave candidata?
  
- Chave primária:
  - chave candidata escolhida no projeto da tabela do banco para identificar unicamente cada registro ou *tupla*.



# Modelo Relacional

- Chave primária:
  - Coluna ou combinação de colunas cujos valores distinguem uma linha ou registro das demais dentro de uma tabela
  - Restrição de chave

chave primária

PROPRIETARIO

CPF	NOME	RUA	NUMERO	BAIRRO
08940256	JOÃO DA SILVA	SAO JOAO	180	CENTRO
03727298	HENRIQUE CARDOSO	IMIGRANTE	1700	VILA 12
97260089	JOSÉ DE SOUZA	SAO JOAO	35	CENTRO



# Modelo Relacional

- Chave primária:
  - Coluna ou combinação de colunas cujos valores distinguem uma linha ou registro das demais dentro de uma tabela
  - Restrição de chave

chave primária composta

				RUA
RUA	TRECHO	NUM_INICIAL	NUM_FINAL	BAIRRO
SAO JOAO	TRC01	0	180	CENTRO
SAO JOAO	TRC02	190	1700	CENTRO
IMIGRANTES	TRC01	0	500	VILA 1



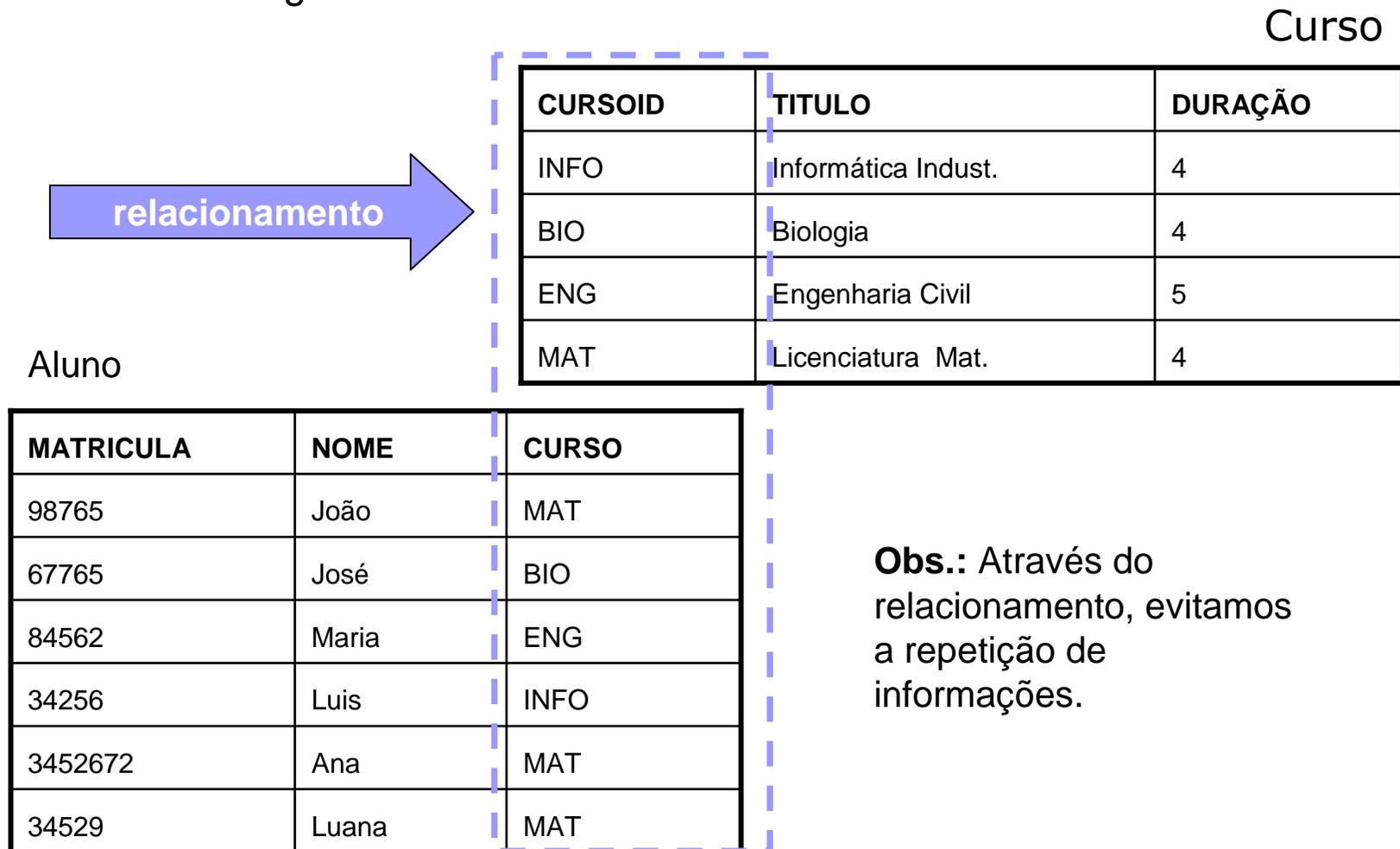
# Modelo Relacional

- Chave estrangeira
  - Implementa a restrição de integridade referencial.
  - Coluna ou combinação de colunas, cujos valores aparecem necessariamente na chave primária de uma outra tabela.
  - Mecanismo que permite a implementação de relacionamentos em um banco de dados relacional.



# Modelo Relacional

- Chave estrangeira





# Modelo Relacional

- Chave estrangeira
  - Uma chave estrangeira não precisa ser uma chave primária na sua relação.
  - Uma chave estrangeira não precisa ter o mesmo nome do que a chave primária correspondente na outra tabela (apenas o mesmo domínio).



# Modelo Relacional

- Chave estrangeira impõe restrições que devem ser garantidas ao serem executadas no BD:
  - Inclusão de uma linha na tabela que contém a chave estrangeira:
    - Garantir que o valor da chave estrangeira exista na chave primária da outra tabela.
  
  - Alteração do valor da chave estrangeira:
    - O novo valor deve aparecer na coluna da chave primária referenciada.



# Modelo Relacional

- Chave estrangeira impõe restrições que devem ser garantidas ao serem executadas no BD:
  - Exclusão de uma linha da tabela que contém a chave primária referenciada por uma chave estrangeira:
    - Não se exclui a linha caso exista um valor na tabela com a chave estrangeira.
    - Remove-se também a linha com o valor de chave estrangeira.
    - Valor da chave estrangeira é ajustado como NULL.



# Modelo Relacional

- Chave estrangeira impõe restrições que devem ser garantidas ao serem executadas no BD:
  - Alteração do valor da chave primária referenciada por alguma chave estrangeira:
    - Propagar a modificação
    - Não deixar que seja feita a modificação



# Álgebra Relacional

- Linguagem de consultas procedural
- Conjunto de operações que usam uma ou duas relações como entrada e geram uma relação de saída
  - *operação*  $(REL_1) \rightarrow REL_2$
  - *operação*  $(REL_1, REL_2) \rightarrow REL_3$
- Operações básicas:
  - Operações unárias:
    - seleção, projeção, renomeação
- Operações binárias:
  - produto cartesiano, união e diferença



# Operadores da Álgebra Relacional

- Seleção:
  - seleciona tuplas que satisfazem um certo predicado ou condição

## Cientes

<i>Nome</i>	<i>Registro</i>
João	1
Maria	2
José	3

a) selecionar tuplas cujo nome = João

$\sigma_{\text{nome}=\text{"João"}}(\text{Clientes})$

<i>Nome</i>	<i>Registro</i>
João	1



# Operadores da Álgebra Relacional

b) selecionar as tuplas de Clientes cujo registro > 1

$\sigma_{\text{registro} > 1}$  (Clientes)

<i>Nome</i>	<i>Registro</i>
Maria	2
José	3

c) selecionar as tuplas de Clientes com registro > 1 e registro < 3

$\sigma_{\text{registro} > 1 \wedge \text{registro} < 3}$  (Clientes)

<i>Nome</i>	<i>Registro</i>
Maria	2



# Operadores da Álgebra Relacional

- Projeção:
  - gera novas relações excluindo alguns atributos
  - exemplo: projete o atributo nome sobre a relação Clientes

$\Pi_{\text{nome}}(\text{Clientes})$

Clientes

<b>Nome</b>	<b>Registro</b>
João	1
Maria	2
José	3



<b>Nome</b>
João
Maria
José



# Operadores da Álgebra Relacional

- União:
  - união de atributos do mesmo domínio que estão em relações diferentes
  - as relações devem possuir o mesmo número de atributos
  - exemplo: encontre todos os clientes da agência que possuem conta corrente ou empréstimo.
    - Relações existentes na agência:
      - ContaCorrente e Emprestimo



# Operadores da Álgebra Relacional

- União:  $\Pi_{\text{nome}}(\text{ContaCorrente}) \cup \Pi_{\text{nome}}(\text{Emprestimo})$

ContaCorrente

<b>Nome</b>	<b>Conta</b>
João	1
Maria	2
José	3

Emprestimo

<b>Nome</b>	<b>Empréstimo</b>
Paulo	100
Maria	200
Carlos	300

Resultado da união

=

<b>Nome</b>
João
Maria
José
Paulo
Carlos



# Operadores da Álgebra Relacional

- Diferença:
  - tuplas que se encontram em uma relação, mas não em outra
  - exemplo: encontre todos clientes sem empréstimo

$$\Pi_{\text{nome}}(\text{ContaCorrente}) - \Pi_{\text{nome}}(\text{Emprestimo})$$

ContaCorrente

<b>Nome</b>	<b>Conta</b>
João	1
Maria	2
José	3

Emprestimo

<b>Nome</b>	<b>Empréstimo</b>
Paulo	100
Maria	200
Carlos	300

Resultado da diferença

<b>Nome</b>
João
José



# Operadores da Álgebra Relacional

## ■ Produto Cartesiano

- Faz todas as combinações entre as tuplas de duas relações
- Gera uma nova relação formada pela união dessas combinações
- Exemplo: produto cartesiano entre os clientes e os empréstimos de Maria

$\sigma_{\text{emprestimo.nome} = \text{“Maria”}}$  (ContaCorrente X Empréstimo)

<b><i>Nome<sub>cc</sub></i></b>	<b><i>Conta</i></b>	<b><i>Nome<sub>emp</sub></i></b>	<b><i>Empréstimo</i></b>
João	1	Maria	200
Maria	2	Maria	200
José	3	Maria	200



# Operadores da Álgebra Relacional

- Operadores derivados:

- Intersecção

- Selecione tudo que está em ambas relações
- Exemplo: todos os clientes que possuem empréstimo

$\Pi_{\text{nome}}(\text{ContaCorrente})$

$\Pi_{\text{nome}}(\text{Emprestimo})$

ContaCorrente

<b>Nome</b>	<b>Conta</b>
João	1
Maria	2
José	3

Emprestimo

<b>Nome</b>	<b>Empréstimo</b>
Paulo	100
Maria	200
Carlos	300

Resultado da intersecção

<b>Nome</b>
Maria



# Operadores da Álgebra Relacional

- Operadores derivados
  - Junção
    - Inclui um produto cartesiano, seguido de uma seleção (pode ter projeção ao final)
    - Exemplo: nomes dos clientes com conta corrente e número de empréstimo:

$$\Pi_{\text{contacorrente.nome, empréstimo.empréstimo}}$$
$$(\sigma_{\text{contacorrente.nome} = \text{empréstimo.nome}} (\text{ContaCorrente} \times \text{Empréstimo}))$$

$$\Pi_{\text{contacorrente.nome, empréstimo.empréstimo}}$$
$$(\text{ContaCorrente} \bowtie \text{Empréstimo}))$$



# Álgebra Relacional - Resumo

