





Unidade 2 – Arquitetura de Sistemas de Banco de Dados e Modelo Relacional

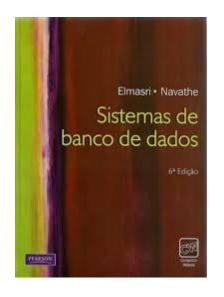


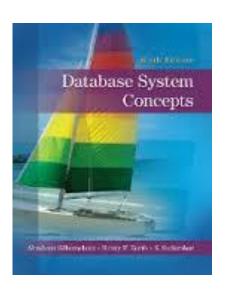
Prof. Aparecido V. de Freitas Doutor em Engenharia de Computação pela EPUSP aparecidovfreitas@gmail.com

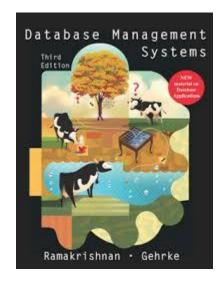


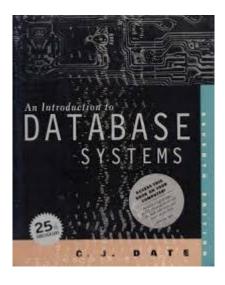


# Bibliografia









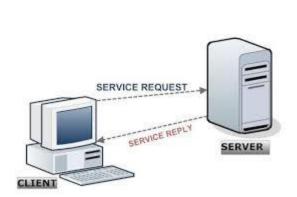




# Introdução



- Primeiros DBMS eram monolíticos (Mainframe).
- Os mais modernos têm um projeto modular, com arquitetura cliente/servidor.
- Módulo Cliente se encarrega da interação com o usuário (GUI). Programas de aplicação e interfaces do usuário, em geral, executam no módulo cliente.
- Módulo Server é responsável pelo armazenamento de dados, acesso, pesquisa e outras funções.









# Abstração de dados

- Abstração de dados corresponde à supressão dos detalhes da organização e do armazenamento dos dados.
- Por meio da abstração de dados, destacam-se recursos essenciais para um melhor conhecimento dos dados.
- Por meio desse conceito, diferentes usuários podem perceber os dados em diferentes níveis de detalhe.
- Uma das principais características da abordagem de banco de dados é prover abstração de dados.





# Como o DBMS provê a abstração de dados?







# Modelo de dados

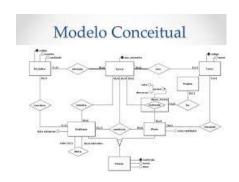
- O modelo de dados é um conceito usado para descrever a <u>estrutura do do banco</u> de dados, oferecendo meios para se obter a abstração de dados.
- A estrutura de um banco de dados inclui os tipos, relacionamentos e restrições que se aplicam aos dados.

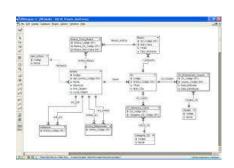




# Categorias de Modelo de dados

- ✓ Modelos de Alto Nível ou modelos conceituais. (Mais próximos da forma pela qual os usuários percebem os dados).
- ✓ Modelos de Baixo Nível ou modelos físicos. (Descrevem os detalhes de como os dados são armazenados no meio físico)
- ✓ Entre esses dois extremos está uma classe de modelos de implementação. (Ocultam muitos detalhes de armazenamento de dados e estão associados ao SGBD particular, ex. Relacional -> tabelas, Hierárquico -> Árvore)



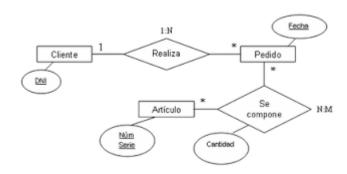






# Modelo Conceitual

- Popularmente definido pelo Modelo Entidade-Relacionamento.
- Utilizam os conceitos de entidades, relacionamentos e atributos.
- Entidades representam objetos do mundo real.
- Atributos representam características dessas entidades.
- Relacionamentos representam associações entre entidades.







# Modelos Físicos

- Descrevem o armazenamento dos dados, com informações de formatos de registros, ordenações de registros e caminhos de acesso.
- Caminhos de acesso tornam eficiente a busca por registros de um banco de dados.
- Îndices são caminhos de acesso usualmente empregados.

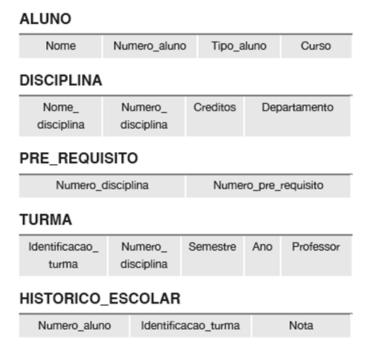






# Esquema

- Um esquema corresponde à descrição do banco de dados, a qual é especificada durante o projeto do banco de dados.
- A representação de um esquema é chamada de diagrama de esquema.







# Instância

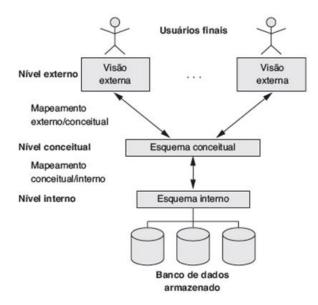
- Os dados reais armazenados em um banco de dados podem mudar com muita frequência.
- Os dados em um determinado instante são chamados de <u>estado</u> do banco de dados ou <u>instância</u> do banco de dados.
- Quando definimos um novo banco de dados, especificamos seu esquema. Nesse ponto, o estado do banco de dados é vazio, sem dados.
- Obtem-se o estado inicial quando o banco de dados é populado ou carregado com dados iniciais.





# Arquitetura de 3 esquemas

- ✓ **Nível interno**: Descreve a estrutura do armazenamento físico do banco de dados
- ✓ **Nível conceitual**: Descreve a estrutura do banco de dados inteiro para uma comunidade de usuários
- ✓ **Nível externo ou de visão**: Descreve a parte do banco de dados em que um grupo de usuários em particular está interessado







# Linguagens do DBMS

- Linguagem de definição de dados (DDL)
  - Usada pelo DBA e pelos projetistas de banco de dados
- Linguagem de definição de visão (VDL)
  - Especificar visões do usuário e seus mapeamentos ao esquema conceitual
- Linguagem de manipulação de dados (DML)
  - Especificar formas de se manipular os dados do banco de dados.



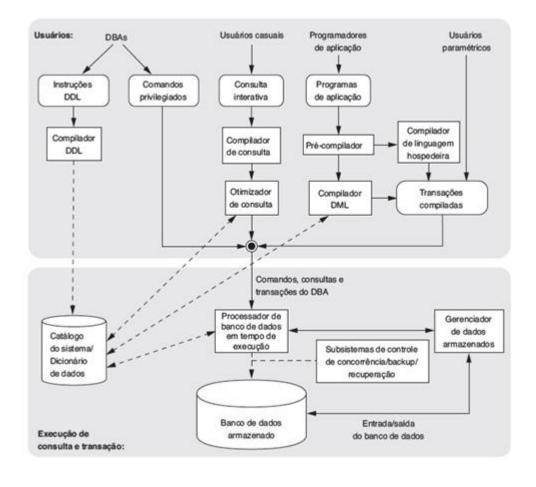






## Ambiente de Sistema de Banco de Dados









## Ambiente de Sistema de Banco de Dados

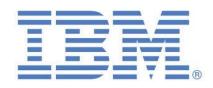
- ✓ Módulos Componentes de um DBMS
  - ✓ Processador de banco em tempo de execução
  - ✓ Catálogo do Sistema
  - ✓ Sistema de Controle de Concorrência
  - ✓ Sistema de Backup e Recuperação
- ✓ Utilitários e ambientes de execução do DBMS
  - ✓ Load (Carga de dados existentes)
  - ✓ Backup (Cópia de segurança dos bancos de dados)
  - ✓ Reorg (Para reorganização dos arquivos de banco de dados)
  - ✓ Monitoração de Desempenho (Coleta de dados para estatística)
  - √ Ferramentas CASE
  - ✓ Ambientes de Desenvolvimento de aplicações

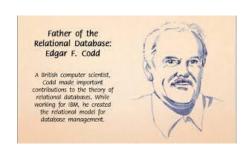




# O Modelo de Dados Relacional











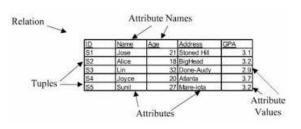
# Modelo Relacional

Desenvolvimento por Tedd Codd, IBM por volta dos anos 70.



- Usa o conceito de relação matemática (parecido com um tabela de valores) e sua base teórica reside na teoria de Conjuntos e Lógica de Predicados de Primeira Ordem.
- As primeiras implementações comerciais foram disponibilizadas a partir de 1980.
- Implantado em uma grande quantidade de sistemas comerciais (SQL/DS, DB2, Oracle, Informix, Sybase, SQL Server, MySQL, PostgreSQL).
- Modelos Hierárquicos e de Rede precederam o modelo relacional.



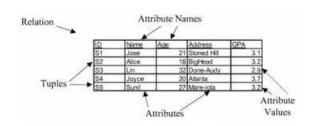






# Modelo Relacional - Conceitos

- No modelo relacional, o banco de dados é representado por um conjunto de relações (informalmente semelhante à uma tabela de valores)
- Cada linha representa uma coleção de valores relacionados.
- Os nomes da tabela e de coluna ajudam a interpretar o significado.







# Tabela - Modelo Relacional

#### ALUNO

Nome	Numero_aluno	Tipo_aluno	Curso
Silva	17.	.1	CC
Braga	8	2	CC

#### DISCIPLINA

Nome_ disciplina	Numero_ disciplina	Creditos	Departamento
Introd. à ciência da computação	CC1310	4	cc
Estruturas de dados	CC3320	4	CC
Matemática discreta	MAT2410	3	MAT
Banco de dados	CC3380	3	cc

#### TURMA

Identificacao_ turma	Numero_ disciplina	Semestre	Ano	Professor
85	MAT2410	Segundo	07	Kleber
92	CC1310	Segundo	07	Anderson
102	CC3320	Primeiro	08	Carlos
112	MAT2410	Segundo	08	Chang
119	CC1310	Segundo	80	Anderson
135	CC3380	Segundo	08	Santos

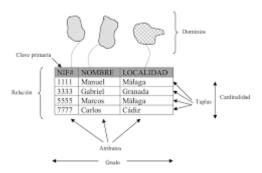
- A primeira tabela chama-se ALUNO pois cada linha representa os fatos de uma entidade particular de aluno.
- Os nomes de coluna Nome, Numero\_Aluno, Tipo\_Aluno e Curso especificam os valores de dados em cada linha.
- Todos os valores em uma coluna são do mesmo tipo.





# Terminologia Formal

- Uma linha em uma tabela é chamada de TUPLA.
- Um cabeçalho da coluna é chamado de ATRIBUTO.
- A tabela é chamada de RELAÇÃO.



 O tipo de dado que descreve os tipos de valores que podem aparecer em cada coluna é representado por um DOMÍNIO de valores possíveis.

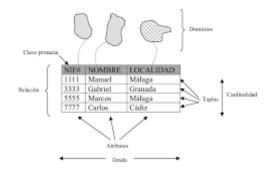


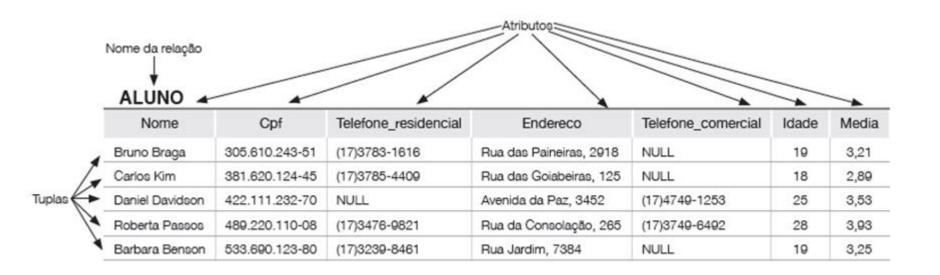




## Domínios, Atributos, Tuplas e Relações

- Domínio: Conjunto de valores atômicos.
- Atômico: Cada valor é indivisível.
- Domínio é especificado por um tipo de dado.









# Esquema Relacional R

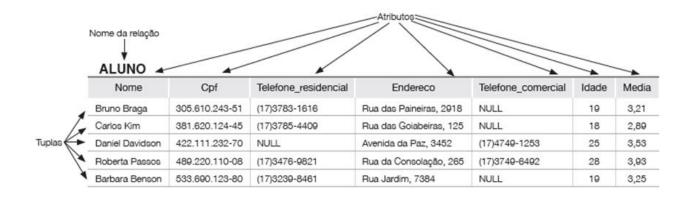
- Um <u>esquema relacional</u> R, indicado por  $R(A_1, A_2, A_3, .... A_n)$  é composto por um nome de relação R e uma lista de atributos:  $A_1, A_2, A_3, ...., A_n$ .
- Os valores de cada atributo A<sub>i</sub>, no <u>esquema relacional</u> R, pertencem ao domínio de A<sub>i</sub>, e é indicado por D(A<sub>i</sub>).
- A descrição de uma relação é feita por meio do <u>esquema relacional</u>.
- O grau (ou aridade) de uma relação é o número de atributos n dessa relação. Por exemplo, uma relação de grau 7, teria sete atributos.





## Relação ou Estado de Relação

- Uma relação ou estado de relação  $\mathbf{r}$  de um esquema relacional  $\mathbf{R}(A_1, A_2, A_3, ..., A_n)$ , também indicado por  $\mathbf{r}(\mathbf{R})$ , é o conjunto de  $\mathbf{n}$  tuplas  $\mathbf{r} = \{ t1, t2, t3, ..., tn \}$ .
- Cada tupla  $t_i$  é uma lista ordenada de n valores  $t_i = \{v_1, v_2, v_3, ...., v_n\}$  no qual cada valor  $v_i$  é um elemento de dom( $A_i$ ) ou é um valor especial chamado NULL.
- Valores NULL representam atributos cujos valores são desconhecidos ou não existem para alguma tupla.







## Definição à luz da Teoria dos Conjuntos

• Uma relação ou estado de relação r, é uma relação matemática de grau n sobre os domínios dom(A<sub>1</sub>), dom(A<sub>2</sub>), dom(A<sub>3</sub>), .... dom(A<sub>n</sub>), que é um subconjunto do produto cartesiano indicado (por X) dos domínios que definem R.:

$$r(R) \subseteq (dom(A_1) \times dom(A_2) \times ... \times dom(A_n))$$

- O produto cartesiano especifica todas as combinações possíveis dos valores dos domínios subjacentes.
- Assim, se indicarmos a CARDINALIDADE (número total de valores em um domínio D) como |D|, considerando que todos os domínios são finitos, o número total de tuplas no produto cartesiano será:

Numero\_Total\_Tuplas(r) = 
$$|dom(A_1)| \times |dom(A_2)| |dom(A_3)| \times ... \times |dom(A_n)|$$





- É possível que vários atributos tenham o mesmo domínio. Ex. Telefone Comercial, Telefone Celular, Telefone Residencial.
- Uma relação é definida como um conjunto de tuplas. Matematicamente, os elementos de um conjunto não possuem ordem entre si. Assim, uma relação não é sensível à ordenação de tuplas. Exemplo: as relações abaixo são idênticas.

### **ALUNO**

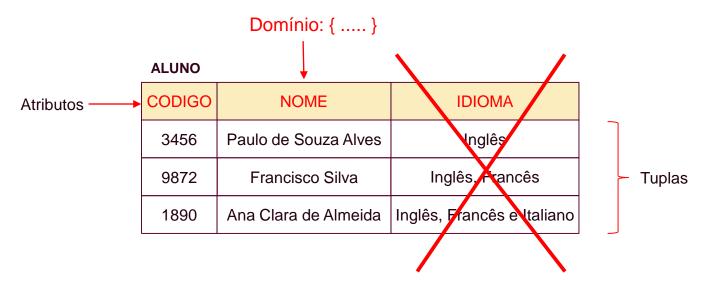
Nome	Cpf	Telefone_ residencial	Endereco	Telefone_ comercial	Idade	Media
Daniel Davidson	422.111.232-70	NULL	Avenida da Paz, 3452	(17)4749-1253	25	3,53
Barbara Benson	533.690.123-80	(17)3239-8461	Rua Jardim, 7384	NULL	19	3,25
Roberta Passos	489.220.110-08	(17)3476-9821	Rua da Consolação, 265	(17)3749-6492	28	3,93
Carlos Kim	381.620.124-45	(17)3785-4409	Rua das Goiabeiras, 125	NULL	18	2,89
Bruno Braga	305.610.243-51	(17)3783-1616	Rua das Paineiras, 2918	NULL	19	3,21

Nome	Cpf	Telefone_residencial	Endereco	Telefone_comercial	Idade	Media
Bruno Braga	305.610.243-51	(17)3783-1616	Rua das Paineiras, 2918	NULL	19	3,21
Carlos Kim	381.620.124-45	(17)3785-4409	Rua das Goiabeiras, 125	NULL	18	2,89
Daniel Davidson	422.111.232-70	NULL	Avenida da Paz, 3452	(17)4749-1253	25	3,53
Roberta Passos	489.220.110-08	(17)3476-9821	Rua da Consolação, 265	(17)3749-6492	28	3,93
Barbara Benson	533.690.123-80	(17)3239-8461	Rua Jardim, 7384	NULL	19	3,25





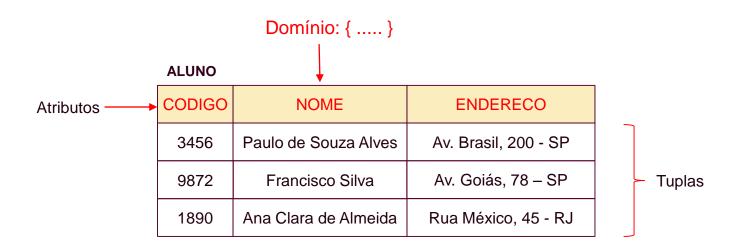
- Cada valor em uma tupla é um valor atômico. Assim, atributos multivalorados não são permitidos. Esse pressuposto é chamado PRIMEIRA FORMA NORMAL.
- Atributos multivalorados possuem um ou mais valores para o mesmo atributo. Por exemplo, o atributo idioma de uma entidade ALUNO pode conter os valores inglês e francês. Para um outro aluno poderia conter apenas um valor - espanhol. Para um terceiro aluno, poderíamos ter 3 valores para este atributo.
- Atributos multivalorados são representados por relações separadas.







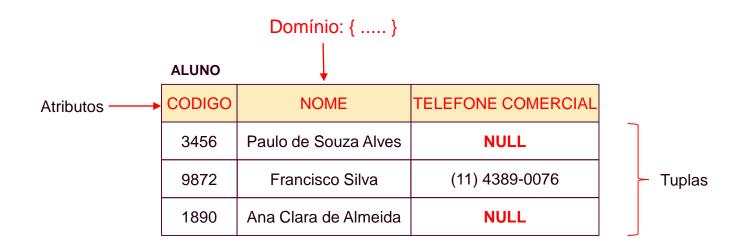
- Atributos compostos são atributos que podem ser divididos em partes menores, ou subpartes, os quais representariam atributos básicos mais simples com significados independentes.
- Por exemplo, um atributo ENDERECO pode ser subdividido em rua, cidade, estado.
  Atributos deste tipo formam uma hierarquia.







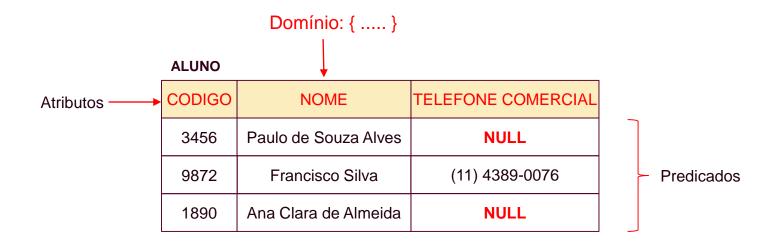
- Valores NULL representam valores que não se aplicam à uma tupla ou são desconhecidos.
- Comparação de valores NULL pode levar à ambiguidades. Por exemplo, se os clientes A e B têm endereços NULL, isso não significa que eles têm o mesmo endereço. Durante o projeto de banco de dados, é conveniente evitar-se o uso de valores NULL.







- Uma interpretação alternativa de uma relação ou estado de relação é como um predicado.
- Nesse caso, os valores em cada tupla são interpretados como valores que satisfazem o predicado, representando desse modo fatos da vida real.
- Essa interpretação é muito útil no contexto das linguagens lógicas de programação, como PROLOG, pois permite que o modelo relacional seja usado nessas linguagens.

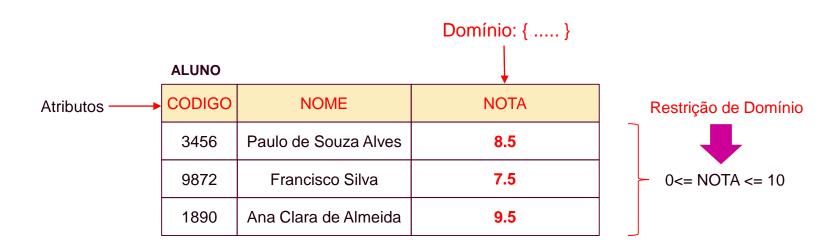






### Restrições de Domínio

Especificam os valores de cada atributo dentro do domínio.







## Restrições de Chaves

- Duas tuplas não podem ter a mesma combinação de valores para todos os seus atributos.
- Ou seja, todos os elementos de um conjunto devem ser distintos.
- No entanto, normalmente existem subconjuntos de atributos de um esquema de relação R, com a propriedade de que duas tuplas em qualquer estado da relação r de R não deverão ter a mesma combinação de valores para esses atributos. Qualquer subconjunto de atributos desse tipo é chamado SUPERCHAVE do esquema de relação R.

	ALUNO			
Atributos	CODIGO	NOME	Curso	CPF
3456		Paulo de Souza Alves	EC	109.098.340-12
	9872	Francisco Silva	CC	239-345-986-12
	1890	Francisco Silva	CC	234.876.456-45





## Superchaves

Considerando o estado da relação abaixo, quais as superchaves da Relação ?

Atributos	CODIGO	NOME	CURSO	CPF
	3456	Paulo de Souza Alves	EC	109.098.340-12
	9872	Francisco Silva	СС	239-345-986-12
	1890	Francisco Silva	СС	234.876.456-45







## Superchave Minima

- Uma superchave pode ter atributos redundantes. Assim, o conceito mais útil é o de uma superchave que não tenha redundância.
- Uma chave de um esquema de relação R é uma superchave de R com a propriedade adicional de que a remoção de qualquer atributo A não mais caracteriza uma superchave.
- Portanto, a chave é uma superchave mínima.

Atributos	CODIGO	NOME	CURSO	CPF
	3456	Paulo de Souza Alves	EC	109.098.340-12
	9872	Francisco Silva	СС	239-345-986-12
	1890	Francisco Silva	СС	234.876.456-45





## Chave

Considerando o estado da relação abaixo, quais as chaves da Relação ?

Atributos	CODIGO	NOME	CURSO	CPF
	3456	Paulo de Souza Alves	EC	109.098.340-12
	9872	Francisco Silva	СС	239-345-986-12
	1890	Francisco Silva	СС	234.876.456-45







### Chave Candidata

- Toda chave também é uma superchave, mas não o contrário.
- Em geral, qualquer superchave formada por apenas um atributo também é chave.
- Em geral, um esquema de relação pode ter mais de uma chave.
- Nesse caso, cada uma das chaves é chamada CHAVE CANDIDATA.





## Chave Primária e Chave Única

- Na modelagem relacional, designa-se uma das <u>chaves</u> <u>candidatas</u> como <u>chave</u> <u>primária</u> (PK = primary key).
- A escolha da chave primária é arbitrária, porém prefere-se escolher uma chave primária com um único atributo.
- Valores NULL não são permitidos para chaves primárias.
- No esquema de relação, indica-se a PK com sublinhado.
- As outras chaves candidatas são chamadas Chaves Únicas (UK = Unique Key)





## Chave primária e Chave única

Considerando o estado da relação abaixo, quais as chaves poderiam ser chaves primárias e chaves únicas ?

Atributos	CODIGO	NOME	CURSO	CPF
	3456	Paulo de Souza Alves	EC	109.098.340-12
	9872	Francisco Silva	СС	239-345-986-12
	1890	Francisco Silva	СС	234.876.456-45







## Exemplo

### **CARRO**

Placa	Numero_chassi	Marca	Modelo	Ano
Itatiaia ABC-7039	A6935207586	Volkswagen	Gol	02
Itu TVP-3470	B4369668697	Chevrolet	Corsa	05
Santos MPO-2902	X8355447376	Fiat	Uno	01
Itanhaem TFY-6858	C4374268458	Chevrolet	Celta	99
Itatiba RSK-6279	Y8293586758	Renault	Clio	04
Atibaia RSK-6298	U0283657858	Volkswagen	Parati	04

Relação CARRO, com duas chaves candidatas: Placa e Numero\_Chassi





# Esquema de Banco de Dados Relacional

- ✓ Conjunto de esquemas de relação S = {R1, R2, ..., Rm}
- Conjunto de restrições de integridade RI.

FUNCIO	NARIO								
Pnome	Minicial	Unome	Cpf	Datanasc	Endereco	Sexo	Salario	Cpf_supervisor	Dnr
DEPAR	TAMENT	О							
Dnome	Dnumero	Cpf_gerent	e Data	_inicio_gerente	•				
Dnumero	Diocal	_DEP							
PROJE Projnome		ero Projloo	al Dnu	m					
	LHA_EN								
	DENTE								





### Estado do Banco de Dados Relacional

- ✓ Conjunto de estados de relação DB = {r1, r2, ..., rm}
- Cada ri é um estado de Ri , tal que os estados de relação ri satisfazem as restrições de integridade especificadas em RI
- Estado Válido: Satisfaz todas as restrições no conjunto de restrições de integridade RI.
- ✓ Estado Inválido: Não obedece à todas as restrições de integridade.





## Estado do Banco de Dados Relacional

#### **FUNCIONARIO**

Pnome	Minicial	Unome	Cpf	Datanasc	Endereco	Sexo	Salario	Cpf_supervisor	Dnr
João	В	Silva	12345678966	09-01-1965	Rua das Flores, 751, São Paulo, SP	м	30.000	33344555587	5
Fernando	T	Wong	33344555587	08-12-1955	Rua da Lapa, 34, São Paulo, SP	м	40.000	8886655576	5
Alice	J	Zelaya	99988777767	19-01-1968	Rua Souza Lima, 35, Curitiba, PR	F	25.000	98765432168	4
Jennifer	s	Souza	98765432168	20-06-1941	Av. Arthur de Lima, 54, Santo André, SP	F	43.000	8886655576	4
Ronaldo	K	Lima	66688444476	15-09-1962	Rua Rebouças, 65, Piracicaba, SP	м	38.000	33344555587	5
Joice	A	Leite	45345345378	31-07-1972	Av. Lucas Obes, 74, São Paulo, SP	F	25.000	33344555587	5
André	V	Pereira.	98798798733	29-03-1969	Rua Timbira, 35, São Paulo, SP	м	25.000	98765432168	4
Jorge	E	Brito	8886655576	10-11-1937	Rua do Horto, 35, São Paulo, SP	м	55.000	NULL	1

#### DEPARTAMENTO

Dnome	Dnumero	Cpf_gerente	Data_inicio_gerente
Pecquisa	6	33344555587	22-06-1988
Administração	4	98765432168	01-01-1996
Matriz	1	8886655576	19-06-1981

### LOCALIZACAO\_DEP

Dnumero	Diocal
1	São Paulo
4	Mauá
5	Santo André
5	ltu
5	São Paulo

#### TRABALHA\_EM

Fcpf	Pnr	Horas
12345678966	1	32,5
12345678966	2	7,5
66688444476	3	40,0
45345345376	1	20,0
45345345376	2	20,0
33344555587	2	10,0
33344555587	3	10,0
33344555587	10	10,0
33344555587	20	10,0
99988777767	30	30,0
99988777767	10	10,0
98798798733	10	35,0
98798798733	30	5,0
98765432168	30	20,0
98765432168	20	15,0
88866555576	20	NULL

### **PROJETO**

Projnome	Projnumero	Projlocal	Dnum
ProdutoX	1	Santo André	6
ProdutoY	2	ltu	5
ProdutoZ	3	São Paulo	6
Informatização	10	Mauá	4
Reorganização	20	São Paulo	1
Novosbeneficios	30	Mauá	4

#### DEPENDENTE

Fcpf	Nome_dependente	Sexo	Datanasc	Parentesco
33344555587	Alicia	F	05-04-1988	Filha
33344555587	Tiago	M	25-10-1983	Filho
33344555587	Janaina.	F	03-06-1968	Esposa
98765432168	Antonio	М	28-02-1942	Marido
12345678966	Michael	M	04-01-1988	Filho
12345678966	Alicia	F	30-12-1988	Filha
12345678966	Elizabeth	F	05-05-1967	Esposa





## Restrições de Integridade

✓ Restrição de Integridade de Entidade: Nenhum valor de chave primária pode ser NULL.

<u>Restrição de Integridade Referencial</u>: Especificada entre duas relações. Mantém a consistência entre tuplas nas duas relações.





FK

## Regra de Chave Estrangeira (Foreign Key)

- ✓ A chave estrangeira ocorre quando um atributo de uma relação (FK) for chave primária em outra relação (PK).
- ✓ Em outras palavras, sempre que houver o relacionamento 1:N entre duas tabelas, a tabela 1 receberá a chave primária e a tabela N receberá a chave estrangeira.
- ✓ Os atributos de FK e PK têm o mesmo domínio. Um valor de FK deve existir em PK, ou ser NULL.

#### **DEPARTAMENTO**

### PK

<u>DEPT</u>	NOME_DEPT
10	COMPRAS
20	ENGENHARIA
40	VENDAS
55	FINANCEIRO

#### **FUNCIONARIO**

COD\_FUNC NOME\_FUNC **CPF DEPT** 3456 Paulo de Souza Alves 345.987.123-98 10 Iosé da Silva 9872 987.243.098.01 10 1890 Pedro Rangel de Souza 112.872.340-81 4680 Angela Silva Medeiros 567.982.045-27 40





### Restrições Semânticas

- ✓ Correspondem à regras de negócio.
- ✓ Não são expressas nos esquemas do modelo de dados, e, portanto, são expressas e impostas por programas de aplicação.
- Exemplo: O máximo de horas que um funcionário pode trabalhar em um projeto por semana é 50 horas.
- Exemplo: O salário do funcionário não pode ser superior ao salário do gerente.







## Operações de Mudança de Estado

- ✓ Inserir, Excluir e Alterar (ou Modificar)
- ✓ Se a operação de inserir violar alguma restrição, o DBMS deve rejeitar a inserção.
- ✓ A operação de excluir pode violar a integridade referencial, quando a tupla que está sendo excluída for referenciada por chaves estrangeiras.
- ✓ Pode-se ter a seguinte situação:
  - ✓ Restrict: rejeita a exclusão
  - Cascade: propaga a exclusão, excluindo tuplas que referenciam a que está sendo excluída.
  - ✓ Set Null ou Set Default: modifica os valores de atributo que referenciam a causa da violação.





### Transação



- ✓ Programa em execução.
- ✓ Inclui algumas operações de banco de dados.
- ✓ Garante que o banco de dados permanece em estado consistente após a execução da transação.
- ✓ As operações executadas por uma transação são atômicas, para garantir a consistência do banco de dados.

