



# MODELAGEM DE DADOS

## Aula 9 – Normalização de Dados

**Curso de Ciência da Computação**

Dr. Rodrigo Xavier de Almeida Leão

Cientista de Dados



O clube de futebol quer armazenar as fichas de controle de cada jogo para manter um histórico dos jogos realizados e da atuação dos seus jogadores.

Nessa etapa você precisará responder:

- Quais os possíveis campos que o documento apresenta?
- Podemos identificar alguma tabela no documento?

**JOGO Nº 114**                      **DATA:** 18/06/2018  
**ADVERSÁRIO:** Cacoalense de RO  
**ESTÁDIO:** ARENA DA FLORESTA   **CIDADE:** Rio Branco  
**UF:** AC  
**CAMPEONATO:** Florestal Brasileiro  
**TÉCNICO:** \_\_\_\_\_  
**PLACAR FINAL:** UVFC \_\_\_\_\_ x \_\_\_\_\_ ADVERSÁRIO

**JOGADORES:**

Nº CAMISA	NOME	Nº de Gols na Partida

**PRINCIPAIS EVENTOS DO JOGO:**

TEMPO:	OCORRIDO:
2 min	Gol Nosso Time
3 min	Gol Adversário
15 min	Pênalti Nosso
16 min	Expulsão Time Adversário

**HISTÓRICO DAS ÚLTIMAS GOLEADAS:**

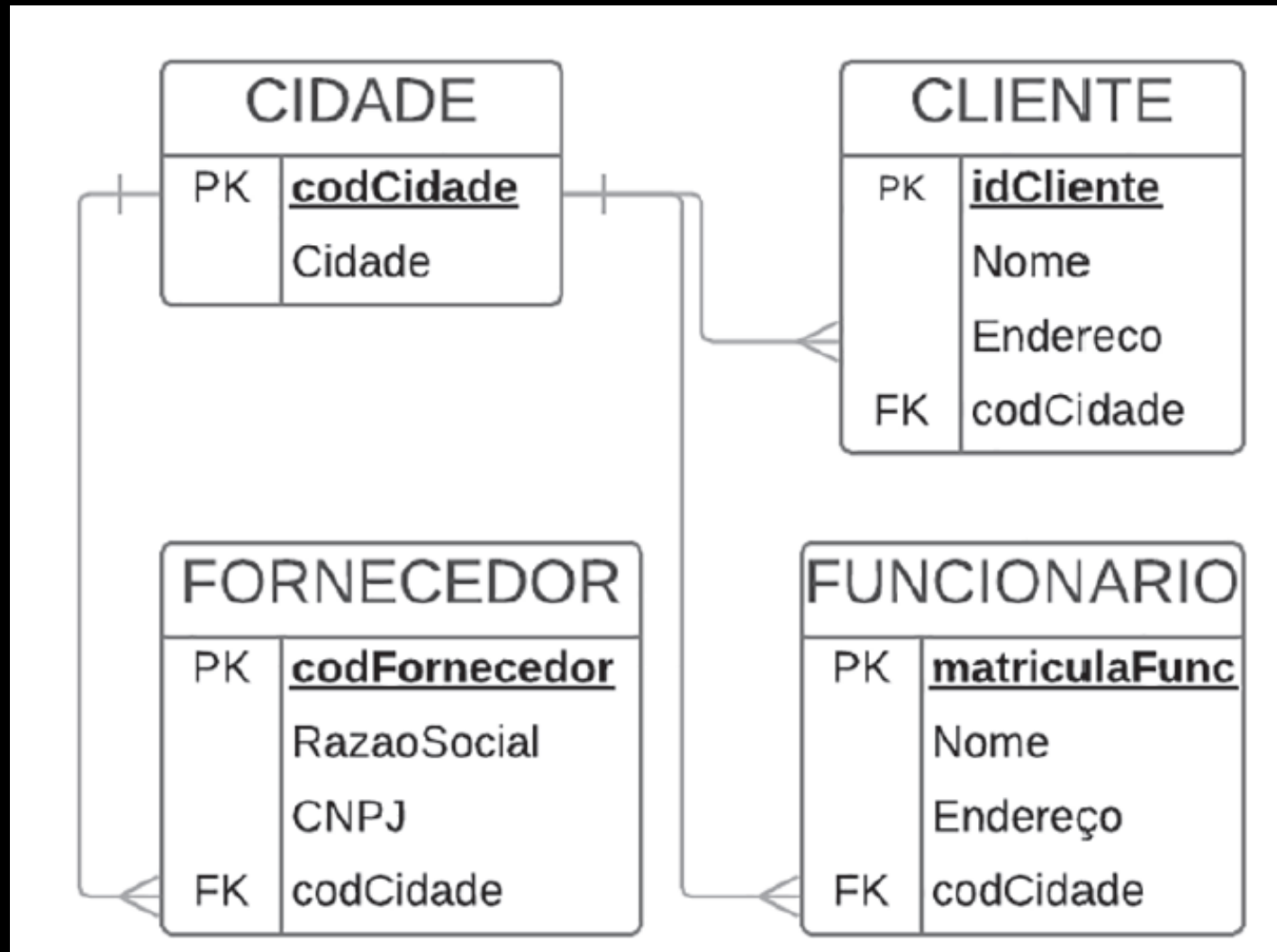
ADVERSÁRIOS	Qtde Gols Adv	Qtde Gols Casa
Pirambu - SE	1	2
Guarabira - PB	0	1
Luverdense - MT	0	2
Social - MG	2	3
Coroatá - MA	2	2

# REDUNDÂNCIA E INCONSISTÊNCIA

CLIENTE		FORNECEDOR		FUNCIONÁRIO	
PK	<u>idCliente</u>	PK	<u>codFornecedor</u>	PK	<u>matriculaFunc</u>
	Nome Endereco CidadeResid		RazaoSocial CNPJ Cidade		Nome CPF CidadeNasc

- **Redundância** ocorre quando uma mesma informação é armazenada mais de uma vez no banco de dados, seja dentro da mesma tabela ou entre diferentes tabelas.
- **Inconsistências:** Diferenças entre as cópias redundantes da informação podem levar a inconsistências nos dados.

# EVITANDO REDUNDÂNCIA E INCONSISTÊNCIA




# \*REDUNDÂNCIA CONTROLADA

O Quadro 4.1 apresenta um outro exemplo de redundância controlada. Observe, na tabela *Funcionário* há o código do departamento (*CodDepartamento*) e o nome do departamento (*NomeDepartamento*), causando a redundância e que pode ser ideal para um determinado software, mas para outro software pode ser uma péssima ideia. É nessa hora que o projetista do banco de dados ou o analista de sistemas deverá decidir se realmente será benéfico deixar a redundância acontecer.

# \*REDUNDÂNCIA CONTROLADA

Tabela: funcionário				
Matricula	Nome	Valor_Hora	CodDepartamento	NomeDepartamento
13467-4	Marco Antonio Liz	R\$ 18,22	DP - 450	Expedição
34562-5	Anna Pietro	R\$ 18,22	DP - 450	Expedição
76321-0	Carlos Werner	R\$ 13,50	DP - 450	Expedição
58309-3	Sandro Lopez	R\$ 28,70	DA - 780	Contabilidade



REDUNDÂNCIA

# NORMALIZAÇÃO

- Normalizar um banco de dados é um processo de projeto que visa organizar as tabelas e seus relacionamentos de forma a reduzir a redundância e garantir a integridade dos dados.
- Para evitar a redundância, é importante normalizar o banco de dados, organizando as tabelas de forma que cada informação seja armazenada em apenas um local seguindo as regras de normalização.
- Isso promove a eficiência, a consistência e a integridade dos dados armazenados.



Podemos listar alguns objetivos e vantagens da normalização de um esquema de banco de dados:

- Diminuição de dados repetitivos deixando o banco de dados mais compacto.
- Aumento da performance no Sistema Gerenciador de Banco de Dados.
- Armazenamento dos dados de forma lógica.
- Facilidade na criação de consultas.
- Permite a concatenação de índices (chaves) de acordo com a quantidade de tabelas envolvidas.
- Facilidade na manutenção do banco de dados.



Quadro 4.2 | Exemplo tabela *Produto* não normalizada

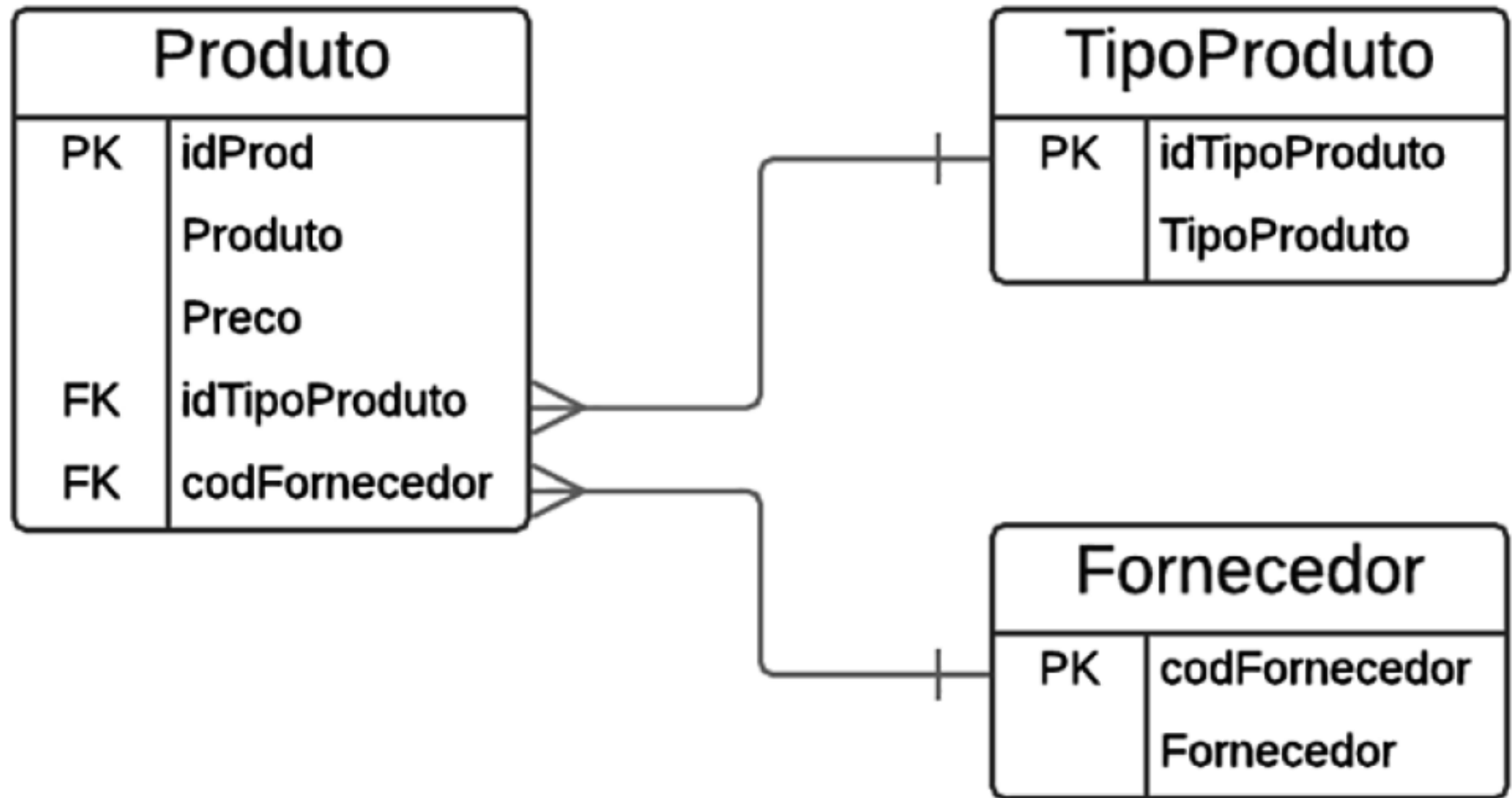
Tabela: produto					
idProd	Produto	Preço	TipoProduto	CodForn	Fornecedor
1415	Sabão	R\$ 4,71	Limpeza	708	Tem Tudo
7841	Álcool	R\$ 5,80	Limpezas	708	Tem de Tudo
8543	Arroz	R\$ 7,84	Grão	516	Compra Boa
9124	Trigo	R\$ 5,45	Grãos	516	Compra B.

Quadro 4.3 | Exemplo tabela *Produto* normalizada

Tabela: produto				
idProd	Produto	Preço	idTipoProduto	CodFornecedor
1415	Sabão	R\$ 4,71	23	708
7841	Álcool	R\$ 5,80	23	708
8543	Arroz	R\$ 7,84	18	516
9124	Trigo	R\$ 5,45	18	516

Tabela: TipoProduto	
idTipoProduto	TipoProduto
23	Limpeza
18	Grãos

Tabela: fornecedor	
codFornecedor	Fornecedor
708	Tem Tudo
516	Compra Boa



# DEPENDÊNCIA FUNICIONAL

Uma dependência funcional em um banco de dados é uma relação entre dois conjuntos de atributos em uma relação (tabela) na qual um conjunto de atributos determina o valor de outro conjunto de atributos. Em outras palavras, uma dependência funcional descreve a relação entre os valores de duas colunas em uma tabela.

Formalmente, uma dependência funcional  $X \rightarrow Y$  significa que, dada uma instância de  $X$ , só pode haver uma instância correspondente de  $Y$ .

->  $X$  é chamado de determinante.

->  $Y$  é chamado de dependente.

ID_Funcionário	Nome	Departamento	Salário
1	João	Vendas	3000
2	Maria	RH	2500
3	José	Vendas	2800
4	Ana	RH	2600

Aqui estão algumas dependências funcionais nesta tabela:

1.  $ID\_Funcionário \rightarrow Nome$

Dado o ID de um funcionário, podemos determinar o nome do funcionário. Cada ID de funcionário está associado a apenas um nome.

2.  $ID\_Funcionário \rightarrow Departamento$

Dado o ID de um funcionário, podemos determinar o departamento em que ele trabalha. Cada ID de funcionário está associado a apenas um departamento.

3.  $Departamento \rightarrow Salário$

Dado o departamento, podemos determinar o salário médio dos funcionários desse departamento. Cada departamento está associado a um salário específico.

Y é dependente funcional de X ou

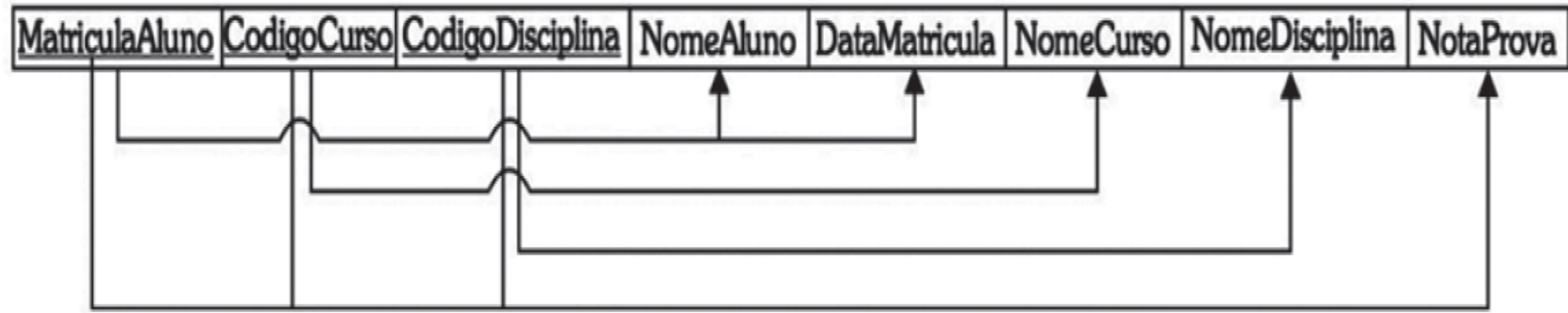
X determina Y ou

Y depende de X, logo

Podemos representar a dependência funcional como:

$$X \rightarrow Y$$

Figura 4.5 | Tabela exemplificando dependência funcional



- Campo *MatriculaAluno*: possui como dependência funcional os campos *NomeAluno* e *DataMatricula*.
- Campo *CodigoCurso*: possui como dependência funcional o campo *NomeCurso*.
- Campo *CodigoDisciplina*: possui como dependência funcional o campo *NomeDisciplina*.
- Campos *MatriculaAluno*, *CodigoCurso*, *CodigoDisciplina*: determinam o valor da *NotaProva*.

#### Quadro 4.6 | Dependência funcional transitiva ou indireta

Tabela: aluno			
<u>Matrícula</u>	Nome	Escola de Origem	Endereço da Escola Origem
1407	Lucca Lews	E.B. Amigos dos Estudos	R. das Montanhas, 450.
5789	Karyn Cruz	E.B. Estudar é Preciso	R. Ventos Fortes, 715.
1587	Jane Flores	E.B. Futuro Melhor	R. Pardal Solitário, 957.

No Quadro 4.6, observe a tabela *Aluno* para um sistema de uma universidade. Temos a necessidade de guardar a escola de origem do aluno e o endereço dessa escola. O endereço da escola de origem é dependente da escola de origem, que depende da chave primária que é a matrícula do aluno. Esse é um exemplo de tabela que precisará sofrer um processo de normalização para resolver essa dependência entre os campos com esse tipo de dependência.



#### Quadro 4.8 | Dependência funcional parcial

Tabela: medição da temperatura			
<u>UF</u>	<u>Cidade</u>	<u>Região</u>	Temperatura
SC	Urubici	Sul	10º
SP	São Carlos	Sudeste	28º
RN	Natal	Nordeste	35º

A dependência funcional parcial ocorre quando um campo ou atributo que não faz parte da chave primária tem dependência funcional de apenas alguns dos atributos que fazem parte da chave primária. No Quadro 4.8 observe que a tabela *Medição da Temperatura*, possui três chaves primárias: UF, Cidade e Região. O campo *Temperatura* possui uma dependência funcional parcial apenas de parte das chaves primárias visto que se o campo *Região* não existisse ou se fosse removido não afetaria o campo *Temperatura*.

#### Quadro 4.7 | Dependência Funcional Total

Tabela: fiscalização		
<u>Cidade</u>	<u>Bairro</u>	Fiscal Responsável
Blumenau	Garcia	Werner Klaus
São Paulo	Ibirapuera	Antônio Luiz
São Paulo	Bom Retiro	Cristina Laís

A dependência funcional total ou completa (como também é conhecida) ocorre quando um atributo que não faz parte da chave primária depende diretamente de todos os outros atributos que fazem parte da chave primária. Sempre ocorre quando a tabela possui chaves concatenadas (mais de uma chave primária). No Quadro 4.7, há um exemplo em que os campos *Cidade* e o *Bairro* são chaves concatenadas e o campo *Fiscal Responsável* depende dos dois campos para “existir”. A dependência funcional total pode ser representada da seguinte forma: *Cidade, Bairro* → *Fiscal Responsável*.