

INSTITUTO FEDERAL
PIAUÍ
Campus Parnaíba

Normalização

Prof. Msc Denival A. dos Santos

Introdução

- Conceito introduzido por Edgar F. Codd, nos em 1970.
- Esta técnica é um processo matemático formal, que tem seus fundamentos na teoria dos conjuntos.
- A **normalização de dados** é uma série de passos que se seguem no projeto de um banco de dados, que permitem um armazenamento consistente e um eficiente acesso aos dados em bancos de dados relacionais. Esses passos reduzem a redundância de dados e as chances dos dados se tornarem inconsistentes, ou seja, objetiva evitar anomalias observadas na inclusão, exclusão e alteração de registros.
- Normalmente após a aplicação das regras de normalização de dados, algumas tabelas acabam sendo divididas em duas ou mais tabelas, o que gera um número maior de tabelas do que o originalmente previsto. Este processo causa a simplificação dos atributos de uma tabela, colaborando significativamente para a estabilidade do modelo de dados, reduzindo-se consideravelmente as necessidades de manutenção.

Problemas de BD não normalizados

- Grupos repetitivos (atributos multivalorados) de dados,
- Atributos compostos,
- Dependências parciais em relação a uma chave concatenada,
- Redundâncias de dados desnecessárias,
- Perdas acidentais de informação,
- Dificuldade na representação de fatos da realidade observada,
- Dependências transitivas entre atributos.

tabela não-normalizada
=
tabela que contém outras tabelas aninhadas

Exemplo de tabela não normalizada

CódProj	Tipo	Descr	Emp					
			CodEmp	Nome	Cat	Sal	DataIni	TempAl
LSC001	Novo Desenv.	Sistema de Estoque	2146	João	A1	4	1/11/91	24
			3145	Sílvio	A2	4	2/10/91	24
			6126	José	B1	9	3/10/92	18
			1214	Carlos	A2	4	4/10/92	18
			8191	Mário	A1	4	1/11/92	12
PAG02	Manutenção	Sistema de RH	8191	Mário	A1	4	1/05/93	12
			4112	João	A2	4	4/01/91	24
			6126	José	B1	9	1/11/92	12

Observação: na coluna **Emp**, aparece, para cada linha de departamento, uma tabela aninhada, que contém os dados dos empregados referentes ao departamento em questão.

Formas normais

- As formas normais a serem vistas serão:
 - Primeira Forma Normal
 - Segunda Forma Normal
 - Terceira Forma Normal
 - Forma Normal Boyce / Codd (FNBC)
 - Quarta Forma Normal
 - Quinta Forma Normal
- **Observação:** vamos estudar as três primeiras que são as principais.

Primeira Forma Normal (1FN)

- A 1FN diz que: cada ocorrência da chave primária deve corresponder a uma e somente uma informação de cada atributo, ou seja, a entidade não deve conter grupos repetitivos (multivalorados).
- Pergunta: **Este campo depende de qual?**
- Passos a serem aplicados para consolidação da 1FN:
 - Identificação da chave primária da tabela.
 - Identificação da coluna que contem dados repetidos e removê-las.
 - Criação de uma nova tabela com chave primária para armazenamento do dado repetido.
 - Criar uma relação entre a tabela principal e a tabela secundária

Primeira Forma Normal (1FN)

Proj:

CódProj	Tipo	Descr
LSC001	Novo Desenv.	Sistema de Estoque
PAG02	Manutenção	Sistema de RH

ProjEmp:

CódProj	CodEmp	Nome	Cat	Sal	DataIni	TempAl
LSC001	2146	João	A1	4	1/11/91	24
LSC001	3145	Sílvia	A2	4	2/10/91	24
LSC001	6126	José	B1	9	3/10/92	18
LSC001	1214	Carlos	A2	4	4/10/92	18
LSC001	8191	Mário	A1	4	1/11/92	12
PAG02	8191	Mário	A1	4	1/05/93	12
PAG02	4112	João	A2	4	4/01/91	24
PAG02	6126	José	B1	9	1/11/92	12

Dependência funcional

- Para descrevermos as próximas formas normais, se faz necessária a introdução do conceito de dependência funcional, sobre o qual a maior parte da teoria de normalização se baseia.
- Dada uma entidade qualquer, dizemos que um atributo ou conjunto de atributos A é dependente funcional de um outro atributo B contido na mesma entidade, se a cada valor de B existir nas linhas da entidade em que aparece, um único valor de A. Em outras palavras, A depende funcionalmente de B.




Dependência funcional

...	Código	...	Salário	...
	E1		10	
	E3		10	
	E1		10	
	E2		5	
	E3		10	
	E2		5	
	E1		10	

- Diz-se que a coluna Salário *depende funcionalmente* da coluna Código (ou que a coluna Código *determina* a coluna Salário) pelo fato de cada valor de Código estar associado sempre ao mesmo valor de Salário. Exemplificando o valor “E1” da coluna Código identifica sempre o mesmo valor de Salário (“10”).
- Para denotar esta dependência funcional, usa-se uma expressão na forma Código → Salário. A expressão denota que a coluna Salário depende funcionalmente da coluna Código. Diz-se que a coluna Código é o *determinante* da dependência funcional.


Dependência Funcional Total (Completa) e Parcial

- Na ocorrência de uma chave primária concatenada, dizemos que um atributo ou conjunto de atributos depende de forma completa ou total desta chave primária concatenada, se e somente se, a cada valor da chave (e não parte dela), está associado um valor para cada atributo.
- Ex: Na entidade ItemPedido, o atributo **QuantProduto** depende de forma total ou completa da chave primária concatenada (**NumPedido+CodProduto**).

ItemPedido	
	CodProduto:
	NumPedido:
	UnidadeProduto:
	QuantProduto: 
	DescProduto:
	ValorUnitario:
	ValorTotal:


Dependência Funcional Total (Completa) e Parcial

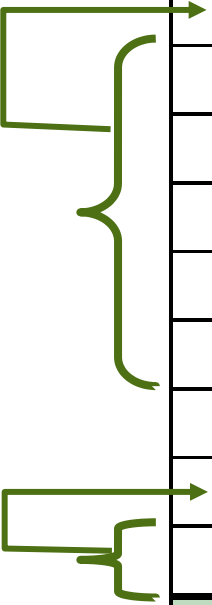
- Na dependência parcial, o atributo só depende de parte da chave primária concatenada e não dela como um todo.
- Ex: Na entidade ItemPedido, o atributo **DescProduto** depende de forma parcial da chave primária concatenada (**NumPedido+CodProduto**), especificamente depende somente da chave primária da entidade ItemPedido (**CodProduto**).

ItemPedido	
	CodProduto:
	NumPedido:
	UnidadeProduto:
	QuantProduto:
	DescProduto: ←
	ValorUnitario:
	ValorTotal:

Dependência Funcional Transitiva

- Quando um atributo ou conjunto de atributos A depende de outro atributo B que não pertence à chave primária, mas é dependente funcional desta, dizemos que A é dependente transitivo de B.
- Ex: Na entidade Pedido, os atributos **Endereço**, **Cidade**, **UF**, **CNPJ** e **InscEstadual** são dependentes transitivos do atributo **NomeCliente**. Nesta mesma entidade, o atributo **NomeVendedor** é dependente transitivo do atributo **CodVendedor**.

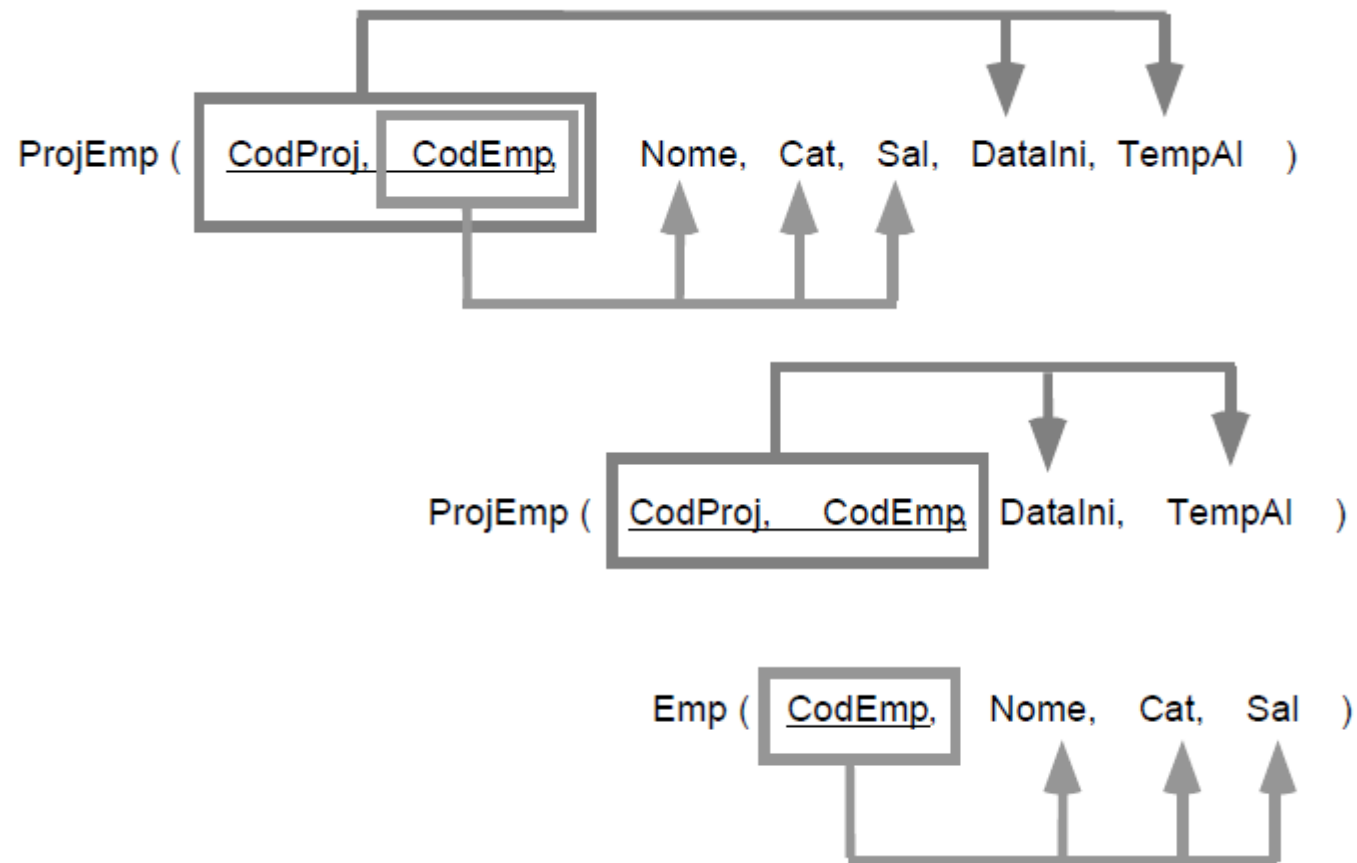
Pedido	
	NumPedido:
	PrazoEntrega:
	NomeCliente:
	Endereço:
	Cidade:
	UF:
	CNPJ:
	InscEstadual:
	TotalPedido:
	CodVendedor:
	NomeVendedor:



Segunda Forma Normal (2FN)

- Uma tabela encontra-se na segunda forma normal (2FN) quando, além de encontrar-se na primeira forma normal, cada coluna não chave depende da chave primária *completa*. Uma tabela que não se encontra na segunda forma normal contém *dependências funcionais parciais*, ou seja, contém colunas não chave que dependem apenas de uma parte da chave primária.
- Pergunta: **Este campo depende de toda a chave?**
- Passos a serem aplicados para consolidação da 2FN:
 - Identificar colunas que não são funcionalmente dependentes da chave primária da tabela.
 - Remover a coluna da tabela e criar uma nova tabela com esses dados.

2FN



Proj:

CódProj	Tipo	Descr
LSC001	Novo Desenv.	Sistema de Estoque
PAG02	Manutenção	Sistema de RH

ProjEmp:

CódProj	CodEmp	DataIni	TempAI
LSC001	2146	1/11/91	24
LSC001	3145	2/10/91	24
LSC001	6126	3/10/92	18
LSC001	1214	4/10/92	18
LSC001	8191	1/11/92	12
PAG02	8191	1/05/93	12
PAG02	4112	4/01/91	24
PAG02	6126	1/11/92	12

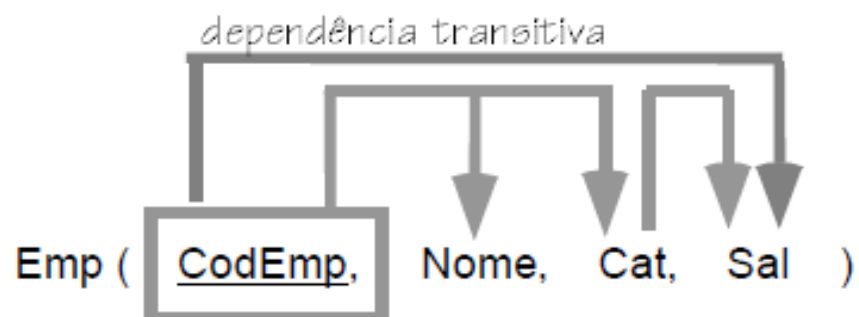
Emp:

CodEmp	Nome	Cat	Sal
2146	João	A1	4
3145	Sílvio	A2	4
6126	José	B1	9
1214	Carlos	A2	4
8191	Mário	A1	4
8191	Mário	A1	4
4112	João	A2	4
6126	José	B1	9

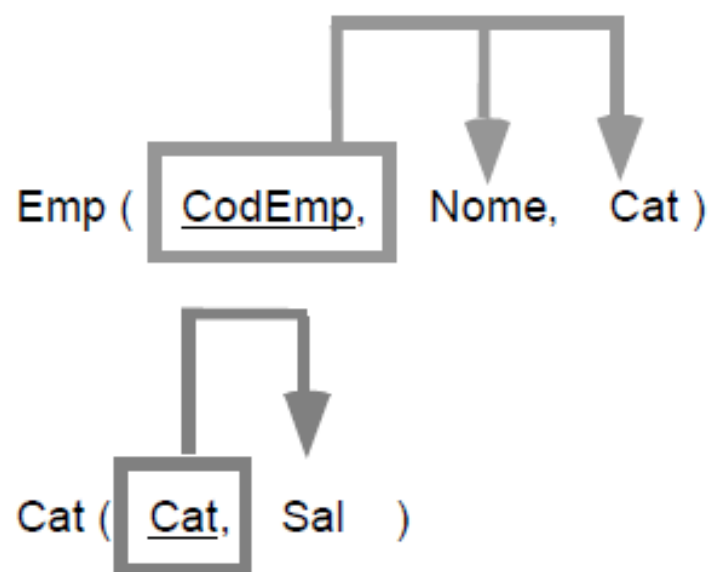
Terceira Forma Normal (3FN)

- Uma tabela encontra-se na 3FN quando, além de estar na 2FN, toda coluna não chave depende *diretamente* de chave primária, isto é, quando não há dependências funcionais *transitivas* ou *indiretas*.
- Pergunta: **Este campo depende de outro que não seja a chave?**
- Passos a serem aplicados para consolidação da 3FN:
 - Identificar as colunas que são funcionalmente dependentes das outras colunas não chave
 - Remover essas colunas

3FN



Tabelas na terceira forma normal (3FN)



Proj:

CódProj	Tipo	Descr
LSC001	Novo Desenv.	Sistema de Estoque
PAG02	Manutenção	Sistema de RH

ProjEmp:

CódProj	NúmEmp	DataIni	TempAl
LSC001	2146	1/11/91	24
LSC001	3145	2/10/91	24
LSC001	6126	3/10/92	18
LSC001	1214	4/10/92	18
LSC001	8191	1/11/92	12
PAG02	8191	1/05/93	12
PAG02	4112	4/01/91	24
PAG02	6126	1/11/92	12

Emp:

NúmEmp	Nome	Cat
2146	João	A1
3145	Sílvio	A2
6126	José	B1
1214	Carlos	A2
8191	Mário	A1
8191	Mário	A1
4112	João	A2
6126	José	B1

Cat:

Cat	Sal
A1	4
A2	4
B1	9

Dados do Pedido

Código do Vendedor: 2541

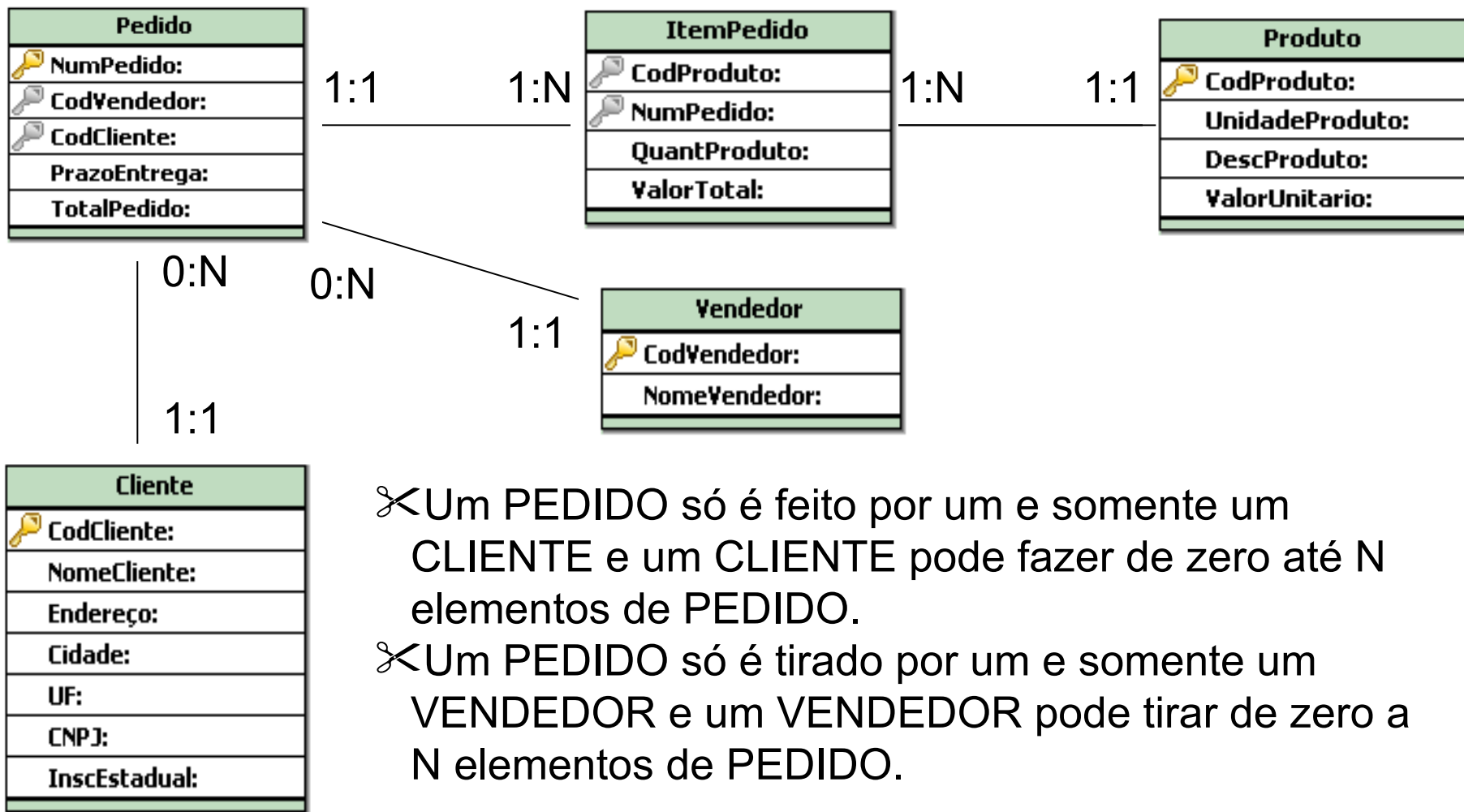
Nome do Vendedor: Gidalto Goes

Prazo de Entrega: 15 dias

num_pedido	Cliente	endereço	CPF	ID_Prod	Produto	Val_Unit	Otde	Unid.	Total	Total Pedido
1854	Elielma Regina	Rua Hugo Machado	78584854854	45	Cartucho	25,00	5	unid	125,00	535,00
1854	Elielma Regina	Rua Hugo Machado	78584854854	100	Papel a4	10,00	10	res	100,00	535,00
1854	Elielma Regina	Rua Hugo Machado	78584854854	350	Grampo	5,00	12	CX	60,00	535,00
3652	Manoel Jailton	Rua da Paz	98568547410	78	DVD	1,00	200	Unid	200,00	535,00
3652	Manoel Jailton	Rua da Paz	98568547410	85	Papel A4	10,00	5	res	50,00	535,00

Exemplo Normalizado

- Após aplicarmos a 3FN teremos o seguinte modelo:



- ✂ Um PEDIDO só é feito por um e somente um CLIENTE e um CLIENTE pode fazer de zero até N elementos de PEDIDO.
- ✂ Um PEDIDO só é tirado por um e somente um VENDEDOR e um VENDEDOR pode tirar de zero a N elementos de PEDIDO.

Exercício - Normalização

IdCD(PK)	NomeCD	Gravadora	Preco	TotalTp	NrFaixa	Musica	Autor	Tempo	Indicacao
1	Mais do Mesmo	EMI	15.00	18:15	1	Será	Renato Russo	2:28	2
1	Mais do Mesmo	EMI	15.00	18:15	2	Ainda é Cedo	Renato Russo	3:55	2
1	Mais do Mesmo	EMI	15.00	18:15	3	Geração Coca-cola	Renato Russo	2:20	2
1	Mais do Mesmo	EMI	15.00	18:15	4	Eduardo e Mônica	Renato Russo	4:32	2
1	Mais do Mesmo	EMI	15.00	18:15	5	Tempo Perdido	Renato Russo	5:00	2
2	Bate-Boca	PolyGram	12.00	14:44	1	Meninos, Eu Vi	Tom Jobim	3:25	1
2	Bate-Boca	PolyGram	12.00	14:44	2	Eu Te Amo	Tom Jobim	3:06	1
2	Bate-Boca	PolyGram	12.00	14:44	3	Piano na Mangueira	Tom Jobim	2:23	1
2	Bate-Boca	PolyGram	12.00	14:44	4	A Violeira	Tom Jobim	2:54	1
2	Bate-Boca	PolyGram	12.00	14:44	5	Anos Dourados	Tom Jobim	2:56	1