

INSTITUTO FEDERAL DE
EDUCAÇÃO, CIÊNCIA E TECNOLOGIA
RIO GRANDE DO NORTE
Campus Mossoró



AULA 25-26

Normalização

Curso: Técnico em Informática (Integrado)

Disciplina: Banco de Dados

Prof. Abrahão Lopes

abrahao.lopes@ifrn.edu.br

Normalização



- “É uma técnica formal para projeto de banco de dados que pode ser usada para realizar uma verificação das estruturas produzidas a partir de diagramas ER.” (GILLENSON, 2006)
- “Normalização é o processo que permite a simplificação da estrutura de um banco de dados de modo que esta se apresente em um estado ótimo, sem duplicação de informações” (DAMAS, 2007)

Normalização



- Ajuda a compreender conceitos importantes como redundância de dados e chaves estrangeiras, e responder perguntas como: que atributos minha entidade deve ter? que relacionamentos devem existir? como fica a cardinalidade?
- Organiza os atributos em tabelas de modo que a redundância seja eliminada, ou seja, remove grupos repetidos de informações.
- Ajuda a melhorar o desempenho e facilitar a manutenção do banco de dados.

Normalização



- Ao projetar um banco de dados deve-se evitar:
 - Existência de grupos repetitivos;
 - Existência de atributos multivalorados;
 - Redundância de dados;
 - Perda de Informações;
 - Falhas na sincronização de dados idênticos em tabelas diferentes;
 - Existência de atributos que dependem apenas de parte de uma chave primária;
 - Existência de dependências transitivas entre atributos;
- A normalização ajuda a resolver esses problemas

Exemplo



- Observe a tabela a seguir:

Produto	Pedido	Cliente	Endereço	Crédito	Data	Vendedor
TV	1546	João	Rua A	6	01-03-10	Marcelo
DVD	2345	Miguel	Av. B	5	12-09-09	Felipe
Rádio	2466	Carlos	Trav. C	9	05-05-10	Marcelo
CD	0987	João	Rua A	11	11-12-08	Antonio
Geladeira	8576	Davi	Rua 9	4	15-04-10	Lidio

Exemplo



- Como as informações sobre produtos, clientes e vendedores estão todas na mesma tabela, existem informações **redundantes** (ex. o endereço de João). Isso gera **desperdício de espaço** e no caso de uma atualização ou consulta **toda a tabela precisa ser percorrida**.
- O ideal é decompor essa tabela em tabelas menores, cada uma armazenando um domínio de dados (produtos, clientes, vendedores, vendas...)
- Antes de estudarmos mais sobre normalização, precisamos compreender sobre **Dependências Funcionais**.

Dependência Funcional



- É uma restrição entre dois conjuntos de atributos de uma mesma entidade/relação representada por uma expressão $X \rightarrow Y$ em que o componente Y é dependente do valor X, ou seja, **existe um valor de X para cada valor de Y**.
- Exemplo:

Produto	Preço	Imposto	Total
TV	400,00	10,00	410,00
DVD	100,00	5,00	105,00
SOM	300,00	8,00	308,00

- Como a coluna **Total** é formada pela soma de Preço e Imposto, dizemos que ela é **dependente funcionalmente** dessas colunas.

Dependência Funcional



- Se X for uma chave primária, todos os demais atributos do registro são funcionalmente dependentes de X.

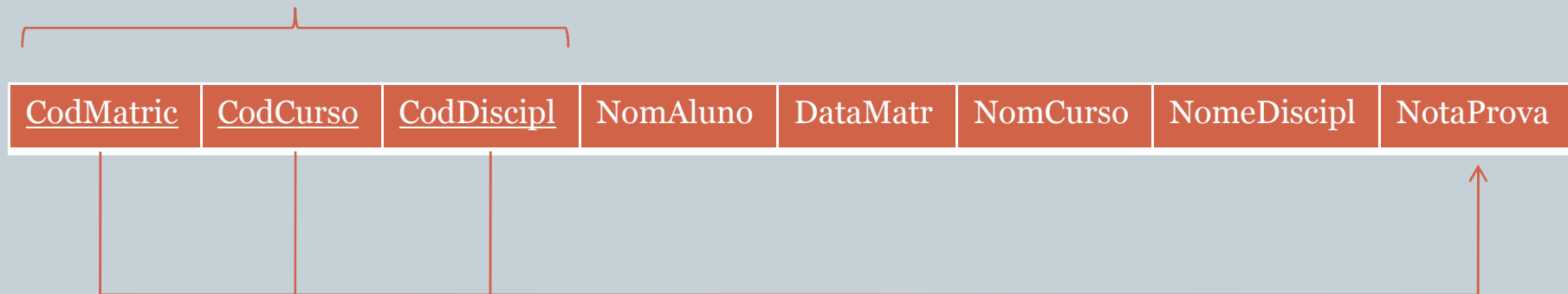
Chave primária → Atributos não chave

- Quando uma chave primária é composta (formada por vários atributos) e um depende de toda a chave, dizemos que há uma **dependência funcional completa (ou total)**.

Dependência Funcional

- Exemplo:

Chave Composta



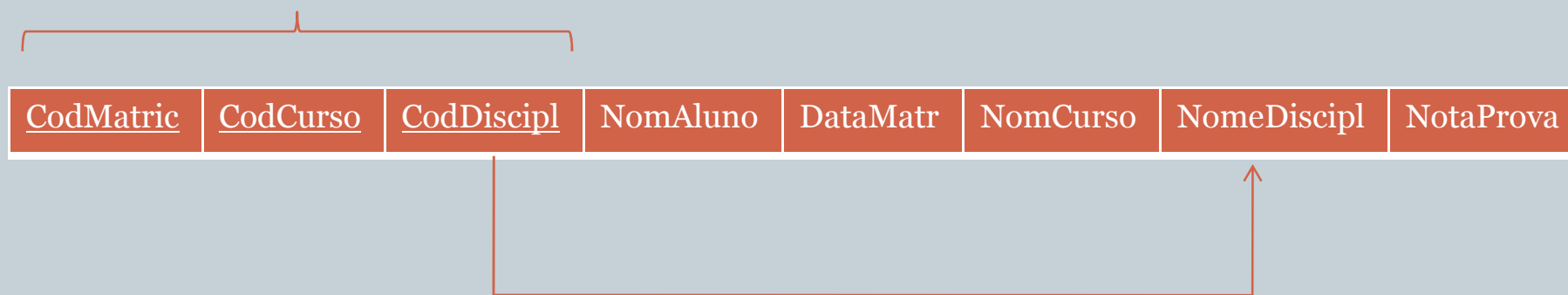
- A nota da prova só poderá ser identificada corretamente quando usamos a chave completa.

Dependência Funcional



- Quando um atributo depende apenas de parte da chave composta, dizemos que há uma **dependência parcial**.

Chave Composta



- O nome da disciplina depende exclusivamente do código da disciplina.

Dependência Funcional



- Se um atributo depende de valores que não é chave ou não fazem parte de uma chave, dizemos que há uma **dependência transitiva**.
- Veremos um exemplo desse tipo de dependência mais adiante.

O processo de normalização



- “A normalização de dados é um processo que nos conduz por uma série de análises para verificação de esquemas de relação quanto à sua situação diante de uma **Forma Normal**.” (ALVES, 2004)
- Existem 5 formas normais, porém as 3 primeiras já oferecem uma boa definição para a estrutura do banco e serão as únicas abordadas aqui.
- Se as regras forem seguidas com cuidado, o sistema será bastante flexível e confiável.

O processo de normalização



- Dizemos que um esquema se encontra em uma determinada forma normal quando **todas as tabelas** do conjunto se encontram normalizadas na mesma forma.

Dados não normalizados (oFN ou FNo)



- Dizemos que os dados não estão normalizados quando todos os atributos se apresentam em um único conjunto.
- Exemplo:

Numero do vendedor, Nome do vendedor, porcentagem da comissão ano de contratação, número do departamento, nome do gerente, número do produto, nome do produto, preço unitário, quantidade.

A primeira forma normal (1FN ou FN1)



- Dizemos que um esquema se encontra na primeira forma normal (1FN) quando **não existem atributos multivalorados**.
- É importante escolher uma **chave primária** para cada tabela.

A primeira forma normal (1FN ou FN1)

- Exemplo de uma tabela com campos multivalorados:

Numer o do vende dor	Numer o do produt o	Nome do vende dor	Porcen tagem de comiss ão	Ano de contra tação	Numer o do depart ament o	Nome do gerent e	Nome do produt o	Preço Unitár io	Quanti dade
137	1000	João	3	1995	20	Carlos	Bola	8,00	4
	1010						Cola	15,00	6
	1030						Goma	2,00	8
221	2000	Marco	5	2000	10	Pedro	Lápis	1,00	9
	3010						Régua	3,00	4
	4020						Pincel	2,00	1
456	3000	Maria	3	1995	20	Carlos	Copo	4,00	2
	9999						Jarro	20,00	6
	0348						Palito	0,50	7

A primeira forma normal (1FN ou FN1)

- Tabela anterior na primeira forma normal:

<u>Numer o do vende dor</u>	<u>Numer o do produt o</u>	Nome do vende dor	Porcen tagem de comiss ão	Ano de contra tação	Numer o do depart ament o	Nome do gerent e	Nome do produt o	Preço Unitár io	Quanti dade
137	1000	João	3	1995	20	Carlos	Bola	8,00	4
137	1010	João	3	1995	20	Carlos	Cola	15,00	6
137	1030	João	3	1995	20	Carlos	Goma	2,00	8
221	2000	Marco	5	2000	10	Pedro	Lápis	1,00	9
221	3010	Marco	5	2000	10	Pedro	Régua	3,00	4
221	4020	Marco	5	2000	10	Pedro	Pincel	2,00	1

- Cada item ocupa sua própria linha (registro).

A primeira forma normal (1FN ou FN1)



- Ou, em modo textual:

Numero do vendedor, Número do produto,
Nome do vendedor, porcentagem da
comissão ano de contratação, número
do departamento, nome do gerente,,
nome do produto, preço unitário,
quantidade.

A primeira forma normal (1FN ou FN1)



- A primeira forma normal é apenas um ponto de partida para a normalização, pois como podemos notar ainda existe muita redundância (ex. o nome do vendedor repetido várias vezes em um mesmo pedido).

<u>Numero do vendedor</u>	<u>Numero do produto</u>	Nome do vendedor
137	1000	João
137	1010	João
137	1030	João

- Isso acontece por que há uma mistura de diferentes domínios de dados numa única tabela (vendedor e produto).

A segunda forma normal (2FN ou FN2)



- Uma tabela encontra-se na segunda forma normal quando, além de estar na primeira forma normal, **cada coluna não chave depende funcionalmente da chave primária completa.** Ou seja, não existem dependências parciais.
- Conseguimos isso decompondo a tabela em **diversas tabelas**, cada uma com seu domínio de dados.

A segunda forma normal (2FN ou FN2)



- Observe que Nome do Vendedor depende apenas de Número do Vendedor e não de Número do Produto.
- O mesmo acontece com Nome do Produto, que depende apenas do Número do produto e não de Número do Vendedor.

A segunda forma normal (2FN ou FN2)



Vendedor (Numero do vendedor, Nome do vendedor, porcentagem da comissão, ano de contratação, número do departamento, nome do gerente) ;

Produto (número do produto, nome do produto, preço unitário).

Quantidade (número do produto, número do vendedor, quantidade).

A segunda forma normal (2FN ou FN2)



Tabela VENDEDOR

<u>Número do vendedor</u>	Nome do vendedor	Porcentagem de comissão	Ano contratação	Número departamento	Nome do gerente
137	João	3	1995	20	Carlos
221	Marco	5	2000	10	Pedro
456	Maria	3	1995	20	Carlos

Tabela PRODUTO

<u>Número do produto</u>	Nome do produto	Preço Unitário
1000	Bola	8,00
1010	Cola	15,00
1030	Goma	2,00

Tabela QUANTIDADE

<u>Número do Vendedor</u>	<u>Número do produto</u>	Quantidade
137	1000	4
137	1010	6
137	1030	8

A segunda forma normal (2FN ou FN2)



- Aqui podemos observar que muita redundância foi eliminada em relação à primeira forma normal. Cada registro aparece apenas uma vez em sua própria tabela.
- Porém nem toda a redundância foi eliminada. Se observarmos, o nome do gerente Carlos aparece duas vezes na tabela vendedor.

<u>Número do vendedor</u>	Nome do vendedor	Porcentagem de comissão	Ano contratação	Número departamento	Nome do gerente
137	João	3	1995	20	Carlos
221	Marco	5	2000	10	Pedro
456	Maria	3	1995	20	Carlos

A segunda forma normal (2FN ou FN2)



- Esse problema acontece por que apesar dos atributos serem dependentes de **Número do vendedor** na relação, o atributo **Gerente** também é funcionalmente dependente de **Número do Departamento**. É o que chamamos de **Dependência Transitiva**.
- Isso nos leva à terceira forma normal.

A terceira forma normal (3FN ou FN3)



- Um esquema está na terceira forma normal quando, além de estar na segunda forma normal, **não existem dependências entre atributos não chave** (dependências transitivas).
- Conseguimos isso decompondo ainda mais as tabelas e criando **chaves estrangeiras**.

A terceira forma normal (3FN ou FN3)



Chave Estrangeira

<u>Número do vendedor</u>	Nome do vendedor	Porcentagem de comissão	Ano contratação	Número departamento
137	João	3	1995	20
221	Marco	5	2000	10
456	Maria	3	1995	20

Número departamento	Nome do gerente
20	Carlos
10	Pedro

A terceira forma normal (3FN ou FN3)



- Nesse ponto nosso banco de dados está livre de redundância, respeitando todas as dependências funcionais e representa a transformação a partir de um diagrama ER.

Exercícios



Referências



- ALVES, W. P. *Fundamentos de Bancos de Dados*. Érica, 2004
- DAMAS, Luiz. *SQL*. 6 ed. LTC, 2007.
- GILLENSON, Mark L. *Fundamentos de Sistemas de Gerência de Banco de Dados*. LTC, 2006.
- HEUSER, Carlos Alberto. *Projeto de Banco de Dados*. Sagra Luzzatto, 2004.
- TEOREY, Toby J. *Projeto e modelagem de banco de dados*. Elsevier, 2007.