

Lanoratório de Banco de Dados

Aula 01

Normalização de Dados

Prof. Me. David Lourenço da Costa

david.costa@metrocamp.edu.br

Normalização de Dados

- “Técnica que visa analisar e refinar arquivos de dados, dividindo-os em vários outros arquivos e estabelecendo relacionamentos entre eles através de informações comuns.” COSTA (2000)
- “Uma forma normal (FN) é uma regra que deve ser obedecida por uma tabela (arquivo) para que seja considerada bem projetada.” HEUSER (2009)

Normalização de Dados

- Objetivos
 - **Evitar** informações repetidas em vários arquivos (redundância) reduzindo o espaço de armazenamento;
 - **Facilitar e agilizar** as atualizações pois serão feitas em um único arquivos;
 - **Garantir** que atualizações nas estruturas de arquivos sejam feitas de maneira mais fácil e segura.

Conceitos Chave

- **Atributo**
- “Atributo é toda e qualquer informação que desejamos registrar através do armazenamento de seu conteúdo em arquivos.” COSTA (2000)
- Características
 - **Tipo** - Determina a natureza da informação que é armazenada como: texto, número, data, imagem, etc.
 - **Tamanho** - Determina o número máximo de caracteres que podem ser armazenados no atributo. Por exemplo, nome pode ser do tipo texto e pode aceitar até 40 caracteres (letras, números ou símbolos como: *, /, %, etc.).

Conceitos Chave

- **Atributo**
- **Características**
 - **Ser obrigatório ou opcional** - Um atributo pode ser obrigatório ou opcional. Por exemplo, o número de matrícula é um atributo obrigatório no arquivo de alunos de uma escola. Já o atributo local de trabalho no mesmo arquivo é opcional, sendo informado somente para os alunos que trabalham e estudam ao mesmo tempo.
 - **Ser condicional a outros atributos** - O preenchimento de um atributo pode estar condicionado a outros atributos. Por exemplo, o atributo cargo no arquivo de alunos está condicionado ao atributo local de trabalho (citado no item anterior). Caso o aluno trabalhe além de informar o local de trabalho, e associado a esse atributo, o cargo exercido deve ser registrado também.

Conceitos Chave

- **Atributo**
- Características
 - **Possuir várias ocorrências** - Um atributo pode ocorrer várias vezes dentro de um mesmo arquivo. Por exemplo, um atributo chamado “séries de repetência” que armazenaria as eventuais séries que o aluno não foi aprovado (repetiu de ano), poderia conter nenhum, um ou mais valores. Um atributo com essa característica também é chamado “atributo multi-valorado”.

Conceitos Chave

- **Registro**

- “Conjunto de atributos que contém informações comuns de um determinado assunto (pessoa, processo, função, objeto, etc.).” COSTA (2000)
- Um exemplo é o registro de funcionários de uma empresa que contém informações (atributos) específicas de cada funcionário como: nome, cargo, salário e departamento que trabalha.

Conceitos Chave

- **Arquivo**

- “Arquivo é um conjunto de registros que tratam do mesmo assunto.” COSTA (2000)
 - Um arquivo pode ser algo físico como um porta CDs ou DVDs, ou eletrônico, como o arquivo de usuários de sites de relacionamento como o facebook ou LinkedIn.

Conceitos Chave

- **Chave**
- “Chave é um atributo, ou um conjunto de atributos, que permite a identificação única (também chamada unívoca) de um registro dentro de um arquivo.” COSTA (2000)
- **Tipos de Chave**
 - **Simples.** É formada por um único atributo. Por exemplo, a chave que identifica um aluno em uma universidade é normalmente formada pelo atributo número de RA (registro acadêmico) do aluno.
 - **Composta.** Formada por mais de um atributo. Por exemplo, a chave que identifica o registro de notas de um aluno deve contar além do RA identificador do aluno, o código da disciplina e o período (semestre e número da prova) correspondente a nota.

Conceitos Chave

- **Dependência Funcional**
- “Dependência funcional ocorre quando um atributo depende de outros atributos para representar uma informação completa.” COSTA (2000)
- Por exemplo, ao analisarmos em um registro de alunos de uma universidade o atributo data de início do curso, esse atributo por si só não transmite uma informação completa, ele nos leva a pergunta: De qual curso ou turma? ou de qual aluno?, portanto, precisamos analisar o atributo **número de RA** para então sabermos que o aluno de RA número X iniciou seu curso em determinada data armazenada no atributo data de início do curso.
- Normalmente os atributos de um registro dependem da chave do registro.

Exemplo



- Observe o Arquivo

<i>Arquivo Vendas</i>	
Data	Data
Cliente	Número
Nome do Cliente	Texto
Endereço do Cliente	Texto
Produto-01	Número
Descrição-01	Texto
Quantidade-01	Número
Valor-01	Número
Produto-02	Número
Descrição-02	Texto
Quantidade-02	Número
Valor-02	Número
Sub-Total	Número
Desconto	Número
Valor Total	Número

- **Resposta:**

- 1) É possível vender mais que 2 produtos para o cliente em uma mesma venda ?;
- 2) Se o endereço do cliente mudar quantos registros devem ser alterados ?;
- 3) Esse arquivo possui informações repetidas ? Se sim, quais ?;
- 4) O valor total é um dado calculado ou informado?

Primeira Forma Normal (1FN)

- Criação de Atributo identificador (Chave)
- Maior clareza em relação a informação do dado

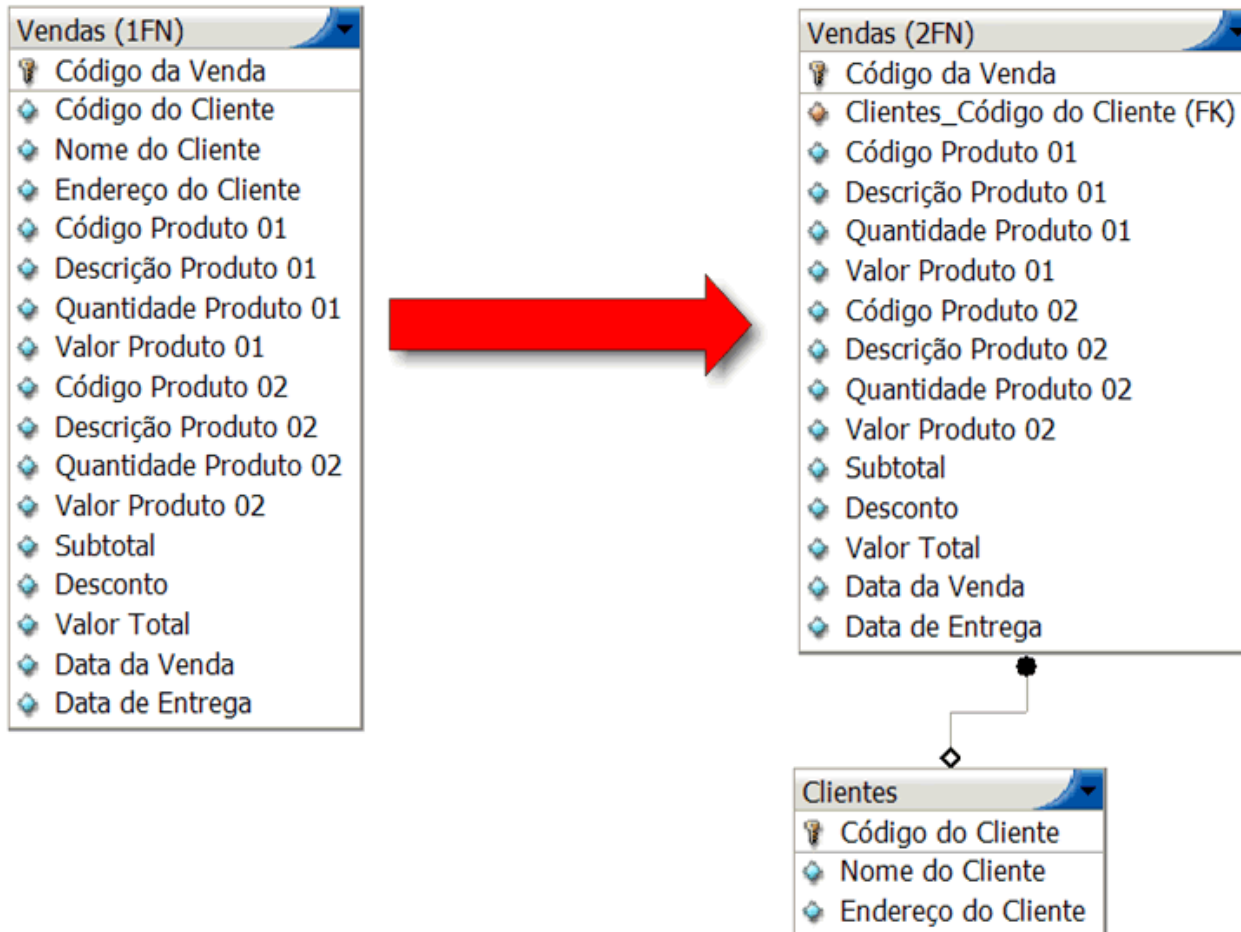
Vendas
◆ Data
◆ Código do Cliente
◆ Nome do Cliente
◆ Endereço do Cliente
◆ Código Produto 01
◆ Descrição Produto 01
◆ Quantidade Produto 01
◆ Valor Produto 01
◆ Código Produto 02
◆ Descrição Produto 02
◆ Quantidade Produto 02
◆ Valor Produto 02
◆ Subtotal
◆ Desconto
◆ Valor Total



Vendas (1FN)
◆ Código da Venda
◆ Código do Cliente
◆ Nome do Cliente
◆ Endereço do Cliente
◆ Código Produto 01
◆ Descrição Produto 01
◆ Quantidade Produto 01
◆ Valor Produto 01
◆ Código Produto 02
◆ Descrição Produto 02
◆ Quantidade Produto 02
◆ Valor Produto 02
◆ Subtotal
◆ Desconto
◆ Valor Total
◆ Data da Venda
◆ Data de Entrega

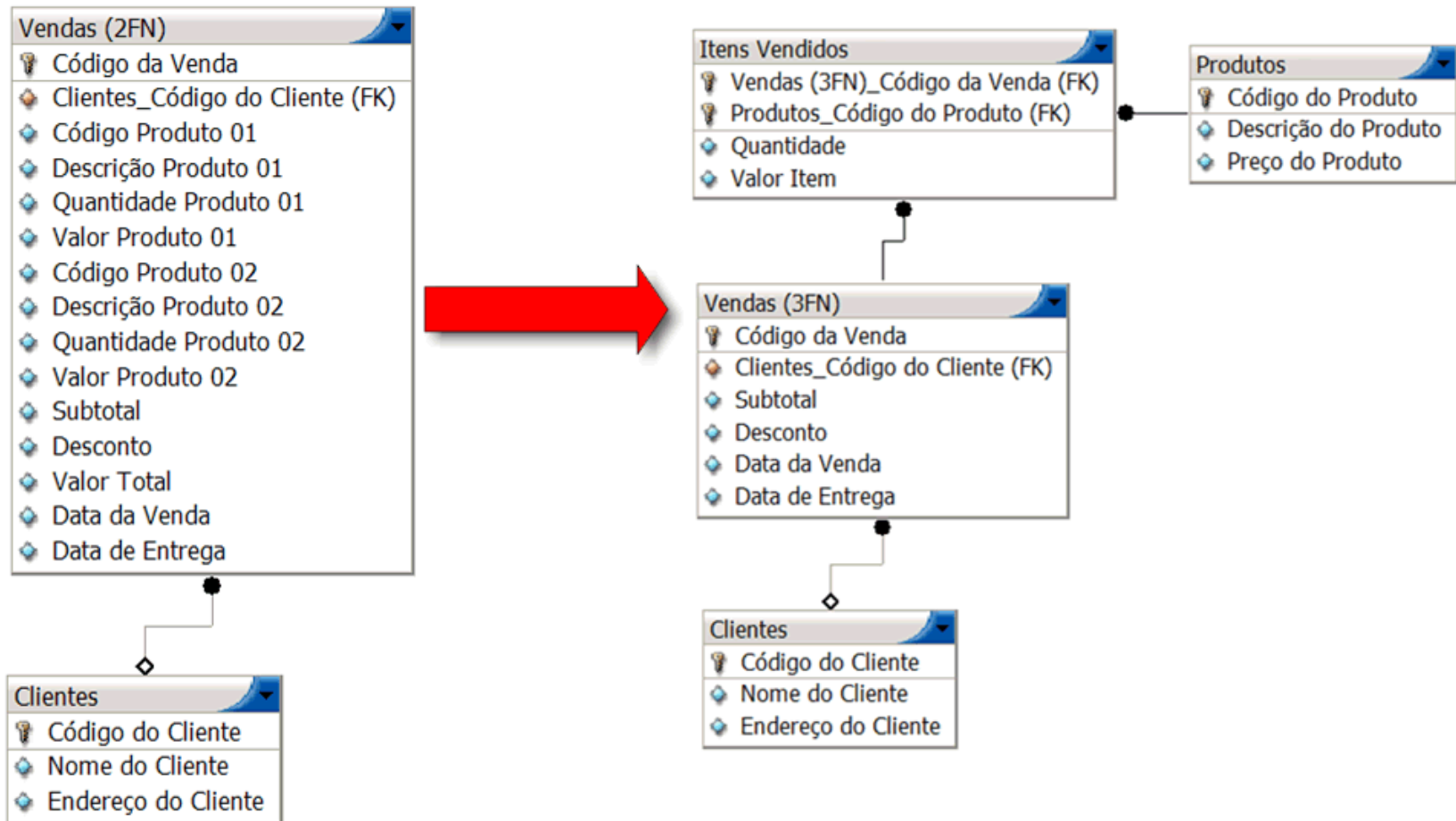
Segunda Forma Normal (2FN)

- Dependência em relação ao atributo identificador



Terceira Forma Normal (3FN)

- Eliminação de atributos multi-valorados e/ou calculados



Desnormalização

- Processo inverso a normalização
- Usado em ambientes OLAP para:
 - Manter histórico de dados voláteis;
 - Minimizar operações de junção (*joins*);
 - Minimizar operações de cálculo usadas em agregações.

Itens Vendidos	
🔑	Código da Venda
🔑	Código do Produto
💎	Quantidade
💎	Valor Item



Itens Vendidos (2)	
🔑	Código da Venda
🔑	Código do Produto
💎	Quantidade
💎	Valor Item
💎	Preço Unitário

Regras Práticas

- **Identificar a chave do arquivo**
 - Isso será importante para identificar dependências funcionais.
- **Separar os atributos que não dependam da chave**
 - Esses atributos são os redundantes e que representam anomalias de atualização e gasto extra de espaço.

Regras Práticas

- **Criar um arquivo para atributos multivalorados**
 - Permite que os atributos que eram limitados a um número n de ocorrências, não tenham mais essa limitação.
- **Eliminar os atributos calculados**
 - Economiza espaço e garante confiabilidade, pois a informação será calculada em tempo de uso.

Atividade 1

- **Estudo de Caso - Clinica Médica Saúde Sempre**

Analise a ficha apresentada em seguida, que está desnormalizada, correspondente a dados coletados por um médico da clinica médica Saúde Sempre em uma consulta. Em seguida normalize-o em etapas (1FN, 2FN) até atingir a **3FN**.

Atividade 1

Ficha Médica			
Número paciente:		Nome:	
Data Nasc:	Sexo:	Convênio:	
Est.Civil:	RG:	Telef:	
Endereço:			
Consultas			
Núm Consulta	Data	Médico	Diagnóstico
Exames			
Núm Consulta	Exame	Data	Resultados

Atividade 2

• Estudo de Caso - Oficina Mecânica do Tião

A oficina mecânica do Tião possui um arquivo onde são registrados os dados dos clientes e serviços prestados, conforme ilustrado na figura abaixo.

Usando as regras práticas para normalização de dados, normalize esse arquivo até atingir a **3FN**.

ARQUIVO DO TIÃO	
◆	NOME DO CLIENTE
◆	ENDEREÇO
◆	TELEFONE
◆	VEÍCULO
◆	ANO
◆	DATA
◆	DIAGNÓSTICO DO PROBLEMA
◆	SERVIÇOS REALIZADOS
◆	VALOR COBRADO