

FIAP GRADUAÇÃO

ANÁLISE E DESENVOLVIMENTO DE SISTEMAS

BUILDING RELATIONAL DATABASE & SQL

Profa. Rita de Cássia Rodrigues



rita@fiap.com.br

Prof. Francisco D. L Abreu



proffrancisco.abreu@fiap.com.br

ANOMALIAS & NORMALIZAÇÃO

- ☐ Introduzir conceitos de modelagem de dados;
- ☐ Identificar os conceitos de Anomalia;
- ☐ Identificar os conceitos de Normalização.

☐ NORMALIZAÇÃO:

- ✓ Anomalias
- ✓ 1ª Forma Normal (1FN)
- ✓ 2ª Forma Normal (2FN)
- ✓ 3ª Forma Normal (3FN)

- Particularidade ou condição do que é anômalo, fora do comum.
- Característica do que desigual ou falta de regularidade de um corpo, um produto, uma matéria, um fenômeno natural, uma coisa qualquer etc: anomalia de caráter genético.
- Aquilo que se desvia da norma, da média geral, irregularidade.

São problemas que ocorrem em bancos de dados não-normalizados, que podem ter inúmeras origens, no qual gera várias inconsistências nos dados. Sendo classificados como:

Anomalias de Modelagem: São repetições desnecessárias de Atributos, Entidades e Relacionamentos;

Anomalias de Implementação: São falhas de Exclusão, Inclusão e Atualização nos dados das Tabelas.

Uma anomalia de inserção ou inclusão acontece quando, ao inserir um dado incompleto ou que possui relação de outro dado indisponível, este pode gerar uma inconsistência no banco de dados.

Por exemplo:

- Incluir um pedido na tabela Pedidos sem o cadastro cliente antes ;
- Incluir a nota de Aluno na tabela Histórico sem o cadastro da Disciplina;
- Incluir dados de Funcionário na tabela, mas não foi definido um atributo auto incremental para o ID;

Uma anomalia de remoção ou exclusão acontece quando:

- Ao remover um registro de uma tabela Principal, os registros relacionados nas tabelas Filhos serão excluídos. (**Remoção Total**)
- Ao remover um registro e uma tabela Principal, os registros relacionados nas tabelas Filhos não são excluídos. (**Remoção Parcial**)

Ambos cenários gera inconsistências no banco de dados. **Qualquer tipo exclusão é uma anomalia**, é necessário pensar no cenário de **Exclusão Lógica**, onde o dado é apenas ocultado através de uma coluna de controle. Por exemplo: Status

Por exemplo:

- Remover os dados do Cliente X da tabela Cliente, os Pedidos serão excluídos devido a dependência relacional.
- Remover os dados do Aluno Y da tabela Aluno, as Notas serão excluídas devido a dependência relacional.

Uma anomalia de atualização ou modificação acontece quando:

- Atualizar um registro de uma tabela Principal, os registros relacionados nas tabelas Filhos não são modificados (**Atualização Parcial**).
- Atualizar um registro de uma tabela deliberadamente, isto é, os dado(s) assumi uma valor diferente do original e se mantém com os falsos (**Atualização Total**).

Ambos cenários gera inconsistências no banco de dados. **É importante ter um backup e/ou registro de transações de dados (log).**

Por exemplo:

- Alterar as propriedades da Chave Primária da tabela Cliente, principalmente, ao configurar o auto incremental de 2 a 2, invés de 1 a 1. Os respectivos códigos relacionais de Cliente à Tabela Pedido precisam ser modificados.

- Ato de normalizar, de tornar normal: normalização do uso de agrotóxicos.
- Retorno a uma situação normal, comum, habitual: normalização de hábitos arcaicos



- O processo de normalização foi proposto por Dr. E. F. Codd **como uma forma de evitar as anomalias de modelagem**. Sendo assim, entende-se que a normalização:
 - Remove a duplicação de dados e, conseqüentemente, minimiza a redundância.
 - É um processo de transformação das relações (tabelas e relacionamentos) em novas relações pela aplicação de quebra de tabelas;
 - Evitar os problemas provocados por falhas no projeto do Banco de Dados, bem como eliminar a “mistura” de assuntos e as correspondentes redundâncias de dados;
 - Se torna econômico e fácil em corrigir as falhas na fase de projeto, do que depois da implementação e utilização do Banco de Dados.

- O processo de normalização foi proposto por Dr. E. F. Codd **como uma forma de evitar as anomalias de modelagem**. Sendo assim, entende-se que a normalização:
 - Remove a duplicação de dados e, conseqüentemente, minimiza a redundância.
 - É um processo de transformação das relações (tabelas e relacionamentos) em novas relações pela aplicação de quebra de tabelas;
 - Evitar os problemas provocados por falhas no projeto do Banco de Dados, bem como eliminar a “mistura” de assuntos e as correspondentes redundâncias de dados;
 - Se torna econômico e fácil em corrigir as falhas na fase de projeto, do que depois da implementação e utilização do Banco de Dados.

- O processo de normalização é **constituído por um conjunto de formas normais**, que definem quando uma tabela está bem estruturada ou não.
- Utilizamos na prática um conjunto **de Cinco Formas Normais (FN)**: 1ª Forma normal (1FN), 2ª Forma Normal (2FN), 3ª Forma normal (3FN), 4ª Forma normal (4FN) e 5ª Forma normal(5FN).
- Após a aplicação das formas normais, algumas tabelas acabam sendo divididas em duas ou mais, o que no final gera um número maior de tabelas originalmente projetada/modelada.

Uma tabela está na Primeira Forma Normal (1ªFN) **se todos os atributos contiverem apenas dados atômicos**, ou seja, cada atributo pode ter apenas um valor por registro (tupla). **Siga um exemplo incorreto logo abaixo:**

MR da Tabela Aluno

Aluno	
 RA: INT	
Nome: VARCHAR	
Telefone: VARCHAR	
	

Tabela Aluno com dados preenchidos

RA	NOME	TELEFONE
111123467	Maria	(11) 4567-1121
111123468	Bianca	(11)98776-1212
111123469	João	(11)997612-1987 (11)5656-0011

Esta tabela não está na 1ªFN porque o atributo **Telefone possui mais de um dado** por registro (tupla). No caso, o RA 11123489 do João Augusto tem dois telefones.

- Este problema ocorre porque misturamos assuntos em uma mesma tabela ou cenário dessa entidade evoluiu para um novo processo de evolução.
- No caso do Telefone é comum que pessoas tenham mais de um número, por exemplo, Residencial, Comercial, Celular Comercial, Celular Pessoal e etc.
- **Nesse exemplo a inconsistência está na super alocação de dados em um único atributo (ou coluna).** Restringindo o tipo do numero, quantos números o aluno tem, além de não alterar ou remover um número

O QUE DEVE SER REALIZADO PARA EVITAR ESSE PROBLEMA?

ANTES DA 1FN

Tabela Aluno com dados preenchidos

RA	NOME	TELEFONE
111123467	Maria	(11) 4567-1121
111123468	Bianca	(11)98776-1212
111123469	João	(11)997612-1987 (11)5656-0011

DEPOIS DA 1FN

Tabela Aluno após a 1FN

RA	NOME
111123467	Maria
111123468	Bianca
111123469	João

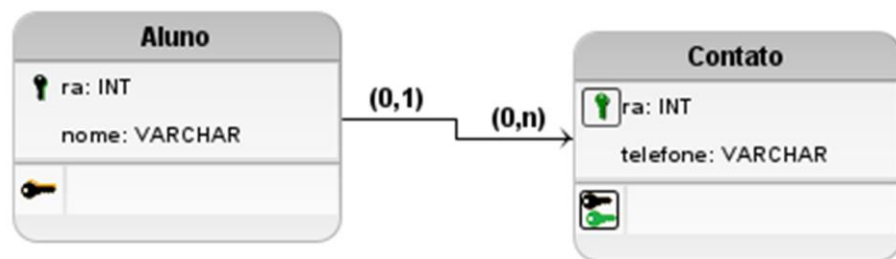
a coluna telefone foi removida

Nova Tabela Contato

RA	TELEFONE
111123467	(11) 4567-1121
111123468	(11)98776-1212
111123469	(11)997612-1987
111123469	(11)5656-0011

a coluna RA é uma chave estrangeira de Aluno

EXPLICATIVO DA SOLUÇÃO



- Criar uma nova tabela (no caso Contato) para conter os dados do atributo não atômico (Telefone);
- Remover o atributo não atômico da tabela original (Aluno).
- Definir uma chave estrangeira para garantir a relação entre a nova tabela e a tabela original;

Uma tabela está na Segunda Forma Normal (2FN) se ela estiver cumprido a 1FN e **que nenhum atributo tenha dependência funcional de qualquer chave candidata**. Siga um exemplo incorreto logo abaixo

MR da Tabela Professor



Professor	
	rm: INT
	nome_prof: VARCHAR
	cpf: VARCHAR
	rg: VARCHAR
	

Tabela Professores com dados preenchidos

RM	NOME_PROF	CPF	RG
10941	FRANCISCO	1111.222.551-11	45.550.12-1
19042	PEDRO	222.551.450-50	33.22.551-x
19042	MARIO	333.550.412.50	148.551.521-52

Esta tabela não está na 2ªFN porque o atributos Nome_Prof tem uma dependência funcional do RG e CPF

- Este problema ocorre porque misturamos assuntos em uma mesma tabela ou cenário dessa entidade evoluiu para um novo processo de evolução.
- No caso, os atributos cpf e rg podem ser considerados uma chave candidata, isto é, uma chave que possui um valor único em algum cenário externo (nível distrital, nacional e internacional).
- **Nesse ultimo exemplo a inconsistência está na dependência entre os atributos dentro da tabela**

O QUE DEVE SER REALIZADO PARA EVITAR ESSE PROBLEMA?

ANTES DA 2FN

Tabela Professores com dados preenchidos

RM	NOME_PROF	CPF	RG
10941	FRANCISCO	1111.222.551-11	45.550.12-1
19042	PEDRO	222.551.450-50	33.22.551-x
19042	MARIO	333.550.412.50	148.551.521-52

DEPOIS DA 2FN

Tabela Professor após a 2FN

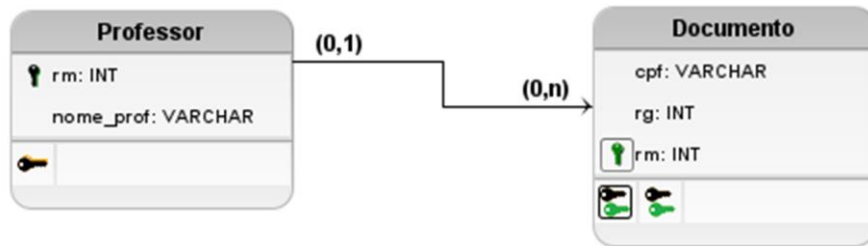
RM	NOME_PROF
10941	FRANCISCO
19042	PEDRO
19042	MARIO

Nova Tabela Documento após a 2FN

RM	CPF	RG
10941	1111.222.551-11	45.550.12-1
19042	222.551.450-50	33.22.551-x
19042	333.550.412.50	148.551.521-52

a coluna rm é uma chave estrangeira de Professor

EXPLICATIVO DA SOLUÇÃO



- Criar uma nova tabela (no caso Documentos) para conter os dados dos atributos dependência parcial (RG e CPF);
- Remover os atributo de dependência parcial da tabela original (Professor).
- Definir uma chave estrangeira para garantir a relação entre a nova tabela e a tabela original;

Uma tabela está na Terceira Forma Normal (3FN) se ela estiver cumprido a 2FN e, **que nenhum atributo tenha dependencia parcial da chave primária e, que nenhum atributo dependa de outro atributo da tabela.** Siga um exemplo incorreto logo abaixo

MR da Tabela Funcionário

Funcionário	
 ID: INT	
ID_LOCAL: INT	
Nome: VARCHAR	
Local: VARCHAR	
	

Tabela Funcionário com dados preenchidos

RM	NOME_FUNC	LOCAL_TRAB	ID_LOCAL
10941	FRANCISCO	Mogi das Cruzes	2
19042	PEDRO	Mogi das Cruzes	4
19042	MARIO	São Paulo	2

Esta tabela não está na 3ªFN porque os atributos Nome_Func depende apenas do RM, enquanto que o atributo Local_Trab depende apenas do ID_Local

- Este problema ocorre porque misturamos assuntos em uma mesma tabela ou cenário dessa entidade evoluiu para um novo processo de evolução.
- No caso, temos campos que não seja dependente diretamente da Chave Primária, ou de parte dela, mas sim dependente de um outro campo constante na tabela, campo este que não seja a Chave Primária
- **Nesse ultimo exemplo a inconsistência está na independência dos atributos comuns da chave primária dentro da tabela**

O QUE DEVE SER REALIZADO PARA EVITAR ESSE PROBLEMA?

ANTES DA 3FN

Tabela Funcionário com dados preenchidos

RM	NOME_FUNC	LOCAL_TRAB	ID_LOCAL
10941	FRANCISCO	Mogi das Cruzes	2
19042	PEDRO	Mogi das Cruzes	4
19042	MARIO	São Paulo	2

DEPOIS DA 3FN

Tabela Funcionário após a 3FN

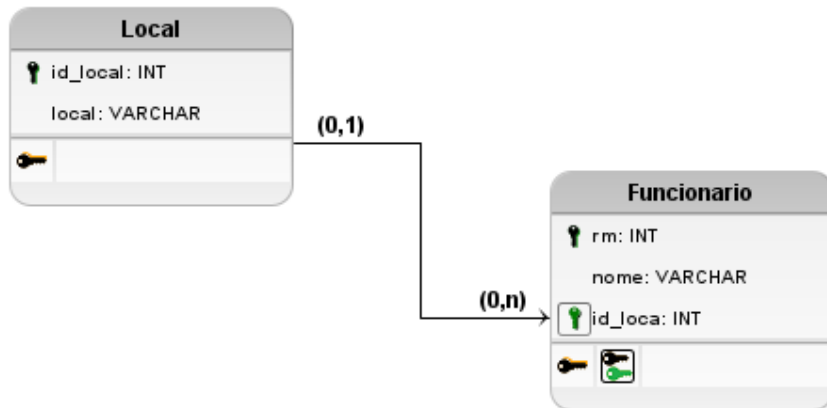
RM	NOME_FUNC	ID_LOCAL
10941	FRANCISCO	2
19042	PEDRO	4
19042	MARIO	2

Nova Tabela Local após a 3FN

LOCAL_TRAB	ID_LOCAL
Mogi das Cruzes	2
Mogi das Cruzes	4
São Paulo	2

a coluna rm é uma chave estrangeira de Professor

EXPLICATIVO DA SOLUÇÃO



- Criar uma nova tabela (no caso Local) para conter os dados dos atributos de independência (`id_local` e `local`);
- Remover os atributo de independência da tabela original (Funcionário).
- Definir uma chave estrangeira para garantir a relação entre a nova tabela e a tabela original;

1FN Não pode ter dados aglomerados

2FN Não pode ter dependência parcial

3FN Não pode ter dependências transitivas



- MACHADO, Felipe Nery R. Banco de Dados - Projeto e Implementação. Érica, 2004. Capítulo 1 – p.19 a 27
- HEUSER, C.A. Projeto de Banco de Dados. Série Livros Didáticos, V. 4. Bookman, 2009. Capítulo 1 –p. 20 a 29
- SILBERSCHATZ, A; KORTH, H. F.; SUDARSHAN, S. Sistema de Banco de Dados. Campus, 2006. Capítulo 6 – p. 133 a 174
- ELMASRI, R.; NAVATHE, S.B. Sistemas de Banco de Dados: Fundamentos e Aplicações. Pearson, 2005. Capítulo 3 – p. 35 a 59

Copyright © 2023 Profa. Rita de Cássia Rodrigues & Prof. Francisco Douglas L. Abreu
Todos direitos reservados. Reprodução ou divulgação total ou parcial deste documento é
expressamente proibido sem o consentimento formal, por escrito, do Professor (autor).