MODELAGEM DE DADOS

Normalização de dados na computação

Em um projeto de banco de dados precisamos lutar contra dois aspectos chaves: redundância de dados e inconsistência de dados. Para lidar com esses fundamentos, nesta webaula, veremos os conceitos sobre normalização de dados e dependência funcional.

CONCEITOS DE NORMALIZAÇÃO DE DADOS

Os danos que a **redundância** pode causar e que geram mais problemas é a repetição da mesma informação em várias tabelas, ocasionando, além de duplicidade, possíveis erros em relatórios. E por que a repetição é a **inconsistência** no banco de dados, que nada mais é do que a mesma informação sendo cadastrada de forma errônea.

A seguir, observe as três tabelas: Cliente, Fornecedor e Funcionário.

CLIENTE	
PK	idCliente
	Nome Endereço CidadeResid

FORNECEDOR	
PK	codFornecedor
	RazaoSocial CNPJ Cidade

FUNCIONÁRIO	
PK	matriculaFunc
	Nome CPF CidadeNasc

Fonte: elaborada pela autora.

Clique aqui e saiba mais

Resolvendo a redundância e a inconsistência entre as tabelas

No caso do campo Cidade é mais simples criamos uma tabela chamada Cidade, relacionando as tabelas que precisam dessa informação.

Fonte: elaborada pela autora.

Redundância controlada

Às vezes, é interessante manter uma informação redundante no banco de dados. Por questões de desempenho de alguma pesquisa ou software, a redundância controlada pode ocorrer. Ela é recomendada quando o campo recebe uma grande quantidade de consultas e poucas alterações. E uma redundância que sabemos que existe, mas é necessária.

Observe que, na tabela Funcionário, há o código do departamento **(CodDepartamento)** e o nome do departamento **(NomeDepartamento)**, fato este que causa a redundância.

Tabela: Funcionário

Matrícula	Nome	Valor_Hora	CodDepartamento	NomeDepartamento
13467-4	Marco Antonio Liz	R$ 18,22	DP - 450	Expedição
34560-5	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

Fonte: elaborada pela autora.

NORMALIZAÇÃO

A **normalização** pode ser definida como um **processo de organizar os dados em um modelo de banco de dados**. Para realizar essas tarefas, são criadas tabelas e estabelecidos relacionamentos entre tabelas de acordo com as regras empregadas, a fim de proteger os dados e tornar o banco de dados mais maleável, adaptando-se para acabar com a redundância e a dependência inconsistente (campo errado na tabela errada) dos dados que serão armazenados no banco de dados.

A normalização de dados é um processo rígido e formal que deve ser seguido passo a passo examinando os campos de uma tabela, com o objetivo de evitar irregularidades observadas na inclusão, exclusão e alteração de registros.

Regras básicas da normalização

Uma das regras básicas da normalização é verificar se determinado campo realmente pertence à tabela.

Tabela Produto não normalizada

Na tabela Produto não normalizada, observe que nos campos TipoProduto e Fornecedor a mesma informação foi inserida erroneamente, entre os registros.

Tabela: Produto

idProd	Produto	Preço	TipoProduto	CodForn	Fornecedor
1415	Sabão	R$ 4,71	Limpeza	708	Tem Tudo
7641	Álcool	R$ 5,80	Limpeza	708	Tem de Tudo
8543	Arroz	R$ 7,84	Grão	516	Compra Boa
9124	Trigo	R$ 5,45	Grãos	516	Compra B.

Fonte: elaborada pela autora.

No caso da tabela de Produtos, há os campos:

- **TipoProduto**: podemos criar uma tabela para armazenar os tipos de produtos.
- **CodForn e Fornecedor**: devemos criar uma tabela chamada Fornecedor.

Tabela Produto normalizada

Veja, a seguir, a normalização com a criação de novas tabelas e campos como chaves estrangeiras. Clique nas abas e veja cada etapa.

TABELA: PRODUTO | TABELA: TIPOPRODUTO | TABELA: FORNECEDOR

Foram retirados da tabela Produto alguns campos, para que esta seja normalizada. As informações que permaneceram e que agora fazem parte de outra tabela ficaram como chaves estrangeiras das tabelas.

Tabela: Produto

idProd	Produto	Preço	idTipoProduto	codFornecedor
1415	Sabão	R$ 4,71	23	708
7641	Álcool	R$ 5,80	23	708
8543	Arroz	R$ 7,84	18	516
9124	Trigo	R$ 5,45	18	516

Fonte: elaborada pela autora.

DER – Produto – TipoProduto - Fornecedor

Diagrama de Entidade-Relacionamento entre as três tabelas:
Produto – Fornecedor e TipoProduto

Fonte: elaborado pela autora.

Geralmente, os analistas de sistemas já criam esses tipos de tabelas relacionadas automaticamente a um modelo de banco de dados. O exemplo é para dar ênfase à necessidade da normalização, principalmente quando estamos começando na modelagem de banco de dados.

TÉCNICAS DE NORMALIZAÇÃO

As técnicas de normalização se baseiam na dependência funcional entre os atributos de uma entidade do banco de dados e nas chaves primárias. A **forma normal** é uma regra que deve ser satisfeita por uma entidade para que ela seja avaliada como uma tabela "projetada com exatidão". São várias formas normais, com regras que vão se tornando mais rigorosas, com o objetivo de averiguar nas tabelas a existência de redundância ou dependências funcionais. No entanto, pelo menos quatro formas normais são consideradas essenciais para a construção de um bom projeto de banco de dados.

Primeira forma normal ou 1FN | Segunda forma normal ou 2FN | Terceira forma normal ou 3FN | Quarta forma normal ou 4FN

DEPENDÊNCIA FUNCIONAL

A dependência funcional consiste em uma restrição entre dois ou mais conjuntos de atributos de uma mesma tabela ou relacionamento (ALVES, 2014).

Um exemplo de dependência funcional é a tabela Aluno.
Nela temos a matrícula e o nome do aluno.
O nome do aluno depende diretamente da matrícula do aluno.

Para melhor entendimento do esquema de relacionamento, explore a galeria a seguir.

Fonte: adaptada de Alves (2014, p. 109).

Os três primeiros campos, MatriculaAluno, CodigoCurso e CodigoDiscipline, representam a chave primária da tabela (no caso, uma chave composta ou concatenada). Neste exemplo o conjunto dos três campos formam um índice único (formando a chave primária).

A dependência funcional pode ser classificada em: transitiva ou indireta, total ou parcial.

Clique em cada um dos tipos para saber mais.

TRANSITIVA OU INDIRETA | TOTAL OU COMPLETA | PARCIAL

Acontece quando um determinado campo da tabela, além de depender da chave primária da tabela, depende de outro campo ou de outros campos que são integrantes da mesma tabela.

Tabela: Aluno

Matrícula	Nome	Escola de Origem	Endereço da Escola Origem
1487	Lucca Lewis	E.B. Amigos dos Estudos	R. Dias Montarmas, 450
5789	Karyn Cruz	E.B. Estudar à Preciso	R. Ventos Fortes, 715
1587	Jane Flores	E.B. Futuro Melhor	R. Pardal Solitário, 957

Fonte: elaborada pela autora.

Exemplo de tabela que precisaria sofrer um processo de normalização para resolver essa dependência entre os campos com esse tipo de dependência.

Normalizar é converter uma tabela em tabelas de graus e cardinalidades menores até que quase não haja redundâncias nem dependências funcionais. Você poderá perceber que o objetivo principal da normalização não é eliminar totalmente as inconsistências, mas sim controlá-las. Identificar as dependências funcionais nas tabelas é o primeiro passo para saber que precisamos normalizar as tabelas em um banco de dados.

Nesta webaula vimos que a normalização de tabelas é um conjunto de procedimentos que permite encontrar erros em um projeto de banco de dados, identificando inconsistências que podem ser informações duplicadas e dependências funcionais mal resolvidas ou mal elaboradas.

Bons estudos!