



Banco de dados com PostgreSQL

Aula 7



Normalização é o processo de modelar o banco de dados projetando a forma como as informações serão armazenadas a fim de eliminar, ou pelo menos minimizar, a redundância no banco. Tal procedimento é feito a partir da identificação de uma anomalia em uma relação, decompondo-as em relações melhor estruturadas.

Um banco de dados dentro dos padrões de normalização reduz o trabalho de manutenção e ajuda a evitar o desperdício do espaço de armazenamento. Se tivermos cadastrado no banco um cliente e tivermos o seu telefone registrado em mais de uma tabela, havendo uma alteração no seu número de telefone, teremos que fazer essa atualização em cada tabela. A tarefa se torna muito mais eficiente se tivermos seu telefone registrado em apenas uma tabela.





Formas Normais

Como mencionado anteriormente, temos conjuntos de regras para determinar com qual forma normal o banco é compatível. Primeiramente, precisamos verificar se encontramos compatibilidade com a primeira forma normal. Caso esteja tudo conforme, analisamos se a segunda forma normal se encaixa e assim sucessivamente.

É importante lembrar que para uma relação atender as exigências de uma forma normal, se faz necessário que esta obedeça as regras da forma normal anterior. A primeira forma normal é exceção pois não existe uma forma normal anterior a primeira.





Primeira Forma Normal

Uma relação está na primeira forma normal quando todos os atributos contém apenas um valor correspondente, singular e não existem grupos de atributos repetidos — ou seja, não admite repetições ou campos que tenham mais que um valor.

O procedimento inicial é identificar a chave primária da tabela. Após, devemos reconhecer o grupo repetitivo e removê-lo da entidade. Em seguida, criamos uma nova tabela com a chave primária da tabela anterior e o grupo repetitivo.





Uma tabela com a FN 1 ainda não aplicada

Código	Nome	Endereço	Telefone
1001	Diego Machado	Rua Tal 321 Porto	5312345678 5398765432
1002	Fulano de Tal	Avenida Tal 71 Centro	5187654321 5143215678





A mesma tabela com a FN 1 aplicada

Código	Nome	Endereço	Bairro
1001	Diego Machado	Rua Tal 321	Porto
1002	Fulano de Tal	Avenida Tal 71	Centro

Código	Telefone
1001	5312345678
1001	5398765432
1002	5112345678
1002	5187654321





Segunda Forma Normal

É dito que uma tabela está na segunda forma normal se ela atende a todos os requisitos da primeira forma normal e se os registros na tabela, que não são chaves, dependam da chave primária em sua totalidade e não apenas parte dela. A segunda forma normal trabalha com essas irregularidades e previne que haja redundância no banco de dados.

Para isso, devemos localizar os valores que dependem parcialmente da chave primária e criar tabelas separadas para conjuntos de valores que se aplicam a vários registros e relacionar estas tabelas com uma chave estrangeira.





Uma tabela com a FN 2 ainda não aplicada

cd_locacao	cd_filme	titulo_filme	devolucao	cd_cliente
1010	201	The Matrix	2011-10-12	743
1011	302	O Grito	2011-12-10	549
1012	201	The Matrix	2011-12-30	362





A mesma tabela com a FN 2 aplicada

cd_filme	titulo_filme
201	The Matrix
302	O Grito

cd_locacao	cd_filme	devolucao	cd_cliente
1010	201	2011-10-12	743
1011	302	2011-12-10	549
1012	201	2011-12-30	362





Terceira Forma Normal

Se analisarmos uma tupla e não encontrarmos um atributo não chave dependente de outro atributo não chave, podemos dizer que a entidade em questão está na terceira forma normal - contanto que esta não vá de encontro as especificações da primeira e da segunda forma normal.

Como procedimento principal para configurar uma entidade que atenda as regras da terceira forma normal, nós identificamos os campos que não dependem da chave primária e dependem de um outro campo não chave. Após, separamos eles para criar uma outra tabela distinta, se necessário.





Uma tabela com a FN 3 ainda não aplicada

placa	modelo	qtd_kmetro	cod_fab	nome_fab
qwe1234	Modelo1	867	3004	fabricante1
asd456	Modelo2	928	3005	fabricante2





A mesma tabela com a FN 3 aplicada

placa	modelo	qtd_kmetro	cod_fab
qwe1234	Modelo1	867	3004
asd456	Modelo2	928	3005

cod_fab	nome_fab
3004	fabricante1
3005	fabricante2





Forma Normal de Boyce-Codd

Uma entidade está na forma normal de Boyce-Codd quando ela estiver na terceira forma normal e não existir dependências multivaloradas entre seus atributos, ou seja, campos que se repetem em relação a chave primária, gerando redundância nas tuplas da entidade. Devemos fragmentar essa relação com o objetivo de não termos mais essas dependências funcionais do gênero.





Uma tabela com a FNBC ainda não aplicada

musica	artista	album
Música 1	Artista 1	Álbum 1
Música 1	Artista 2	Álbum 2
Música 1	Artista 1	Álbum 2
Música 2	Artista 3	Álbum 1
Música 2	Artista 2	Álbum 1
Música 2	Artista 3	Álbum 2





A mesma tabela com a FNBC aplicada

musica	artista
Música 1	Artista 1
Música 1	Artista 2
Música 2	Artista 3
Música 2	Artista 2

musica	album
Música 1	Álbum 1
Música 1	Álbum 2
Música 2	Álbum 1
Música 2	Álbum 2



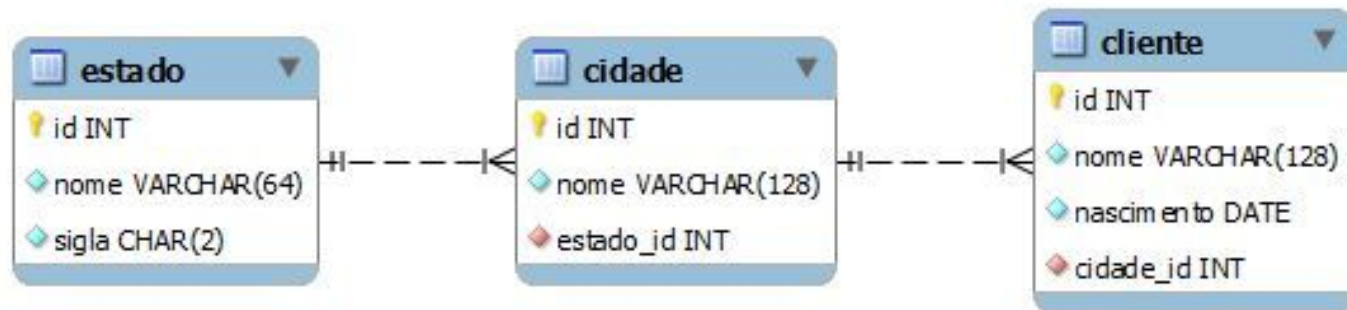


Com o aumento do número de linhas presentes nas tabelas de uma base de dados, percebe-se o aumento da utilização de recursos computacionais para realizar consultas mais complexas, principalmente aquelas que envolvem junções entre duas ou mais tabelas.

É nítido que, mesmo em um cenário ideal, com todas as tabelas normalizadas e perfeitamente indexadas, as junções tendem a ser o mais frequente problema de desempenho nas consultas para geração de relatórios ou análises mais complexas entre os dados dispersos por várias tabelas. Este cenário é ainda mais crítico em aplicações web que disponibilizam dados em tempo real, como a maioria dos sites sociais, onde a informação deve estar atualizada sempre que o usuário executa uma ação ou recarrega uma página.



Uma estrutura já normalizada





A mesma estrutura desnormalizada, possuindo uma redundância controlada.

