



Projeto de Banco de Dados

Prof^a: Flavia Garcia

AULA 3

Normalização de dados

- É uma sequência de etapas sucessivas que, ao final, apresentará um modelo de dados estável com um mínimo de redundâncias e anomalias.

Banco de dados não normalizado

MatFunc	NomeFunc	Dtnasc	EndFunc	Numdepto	NomeDepto	MatGer
F1	Helder	31/08/76	R. do Rio, 72	D1	Logística	F5
F2	Antônio	18/12/85	R. Do Retiro, 10	D2	Qualidade	F6
F3	Bruno	03/07/79	R. Vermelha, 16	D1	Logística	F5
F4	Amorim	17/01/85	R. Monteiro, 25	D2	Qualidade	F6
F5	Artur	23/05/79	R. Paraná, 255	D2	Qualidade	F6
F6	Francisco	28/02/89	R. Acre, 12	D1	Logística	F5

Funcionário_Departamento

Banco de dados não normalizado

Matfunc	Numproj	Horas	NomeFunc	NomeProj	LocalProj
F1	PR3	20	Helder	Serra	Petrópolis, Friburgo
F2	PR1	10	Antônio	Baixada	Caxias, Nova Iguaçu
F3	PR2	30	Bruno	Z. Oeste	Recreio, Barra
F1	PR2	15	Helder	Z. Oeste	Recreio, Barra
F4	PR1	25	Amorim	Baixada	Caxias, Nova Iguaçu
F3	PR3	10	Bruno	Serra	Petrópolis, Friburgo

Funcionário_Projeto

Redundância de dados

- Exemplo:

- Nome do funcionário

- Qualquer modificação no nome do funcionário deve ser realizada em mais de uma tabela;
 - Na tabela Funcionario_Projeto essa alteração deve ser realizada 2 vezes.
 - Qualquer modificação num departamento refletirá em vários campos da tabela Funcionario_Departamento.

Inconsistência de dados

- A cada novo registro de um funcionário que for inserido para o departamento D1 na tabela Funcionario_Projeto, se por acidente for digitado "Lofística" ao invés de "Logística", e numa segunda inserção for digitado "Logóstica", a cada consulta ou a minha consulta deverá conter todos os possíveis erros para a palavra "Logística" ou não retornará o resultado desejado.

Anomalias de exclusão

- Na tabela Funcionario_Projeto, se todos os funcionários alocados em um determinado projeto forem excluídos, serão perdidas também todas as informações referentes a este projeto.
- Exemplo: Se os funcionários Bruno e Helder forem excluídos, não haverá mais nenhuma informação a respeito do projeto Zona Oeste.

Anomalias de modificação

- Ao efetuar uma modificação, esta deve ser replicada manualmente para todas as instâncias daquela informação, além disso, a modificação pode gerar um caso de inconsistência.

Normalização

- A solução para evitar essas anomalias é a normalização formal das nossas informações, tabelas e respectivos relacionamentos.

Normalização de dados

- Há 5 regras que se aplicam a banco de dados:
 - Primeira Forma Normal (1FN)
 - Segunda Forma Normal (2FN)
 - Terceira Forma Normal (3FN)
 - Quarta Forma Normal (4FN)
 - Quinta Forma Normal (5FN)

Normalização de dados

- Ao aplicarmos as regras anteriores já é possível atingirmos o objetivo ao qual se propõe a normalização de dados. Um modelo estável pode ser garantido quando atingimos a 3FN.
- A 4FN e a 5FN são utilizadas em casos específicos, desde que sejam absolutamente necessárias.

Primeira Forma Normal

- Esta relacionada a definição formal da relação: não deve permitir atributos multivalorados, atributos compostos e suas combinações.

Aplicando a 1FN ao exemplo

Matfunc	Numproj	Horas	NomeFunc	NomeProj	LocalProj
F1	PR3	20	Helder	Serra	Petrópolis
F1	PR3	20	Helder	Serra	Friburgo
F2	PR1	10	Antônio	Baixada	Caxias
F2	PR1	10	Antônio	Baixada	Nova Iguaçu
F3	PR2	30	Bruno	Z. Oeste	Recreio
F3	PR2	30	Bruno	Z. Oeste	Barra
F1	PR2	15	Helder	Z. Oeste	Recreio
F1	PR2	15	Helder	Z. Oeste	Barra
F4	PR1	25	Amorim	Baixada	Caxias
F4	PR1	25	Amorim	Baixada	Nova Iguaçu
F3	PR3	10	Bruno	Serra	Petrópolis
F3	PR3	10	Bruno	Serra	Friburgo

Aplicando a 1FN ao exemplo

- Esta primeira solução atende ao requisito de possui somente valores atômicos, porém causa um número grande de redundâncias....
- Uma outra solução seria a definição de uma nova tabela.

Aplicando a 1FN ao exemplo

Matfunc	Numproj	Horas	NomeFunc	NomeProj
F1	PR3	20	Helder	Serra
F2	PR1	10	Antônio	Baixada
F3	PR2	30	Bruno	Z. Oeste
F1	PR2	15	Helder	Z. Oeste
F4	PR1	25	Amorim	Baixada
F3	PR3	10	Bruno	Serra

Cod_local	Numproj	Local_Proj
L1	PR1	Caxias
L2	PR1	Nova Iguaçu
L3	PR2	Barra
L4	PR2	Recreio
L5	PR3	Petrópolis
L6	PR3	Friburgo

Segunda Forma Normal (2FN)

- Tem o objetivo de identificar dependências funcionais parciais e com isso identificar as redundâncias que existam ainda na 1FN.
- Para isso é necessário analisar as tabelas cuja chave primária seja composta e verificar se não existe qualquer campo dependente exclusivamente da chave primária.

Dependência funcional

- Existência de campos em uma tabela cuja a ocorrência de valores está associada a valores que são preenchidos em outros campos nessa mesma tabela.
- Tomando como exemplo a tabela Funcionario_Departamento:

Dependência Funcional

- O campo NomeFunc é dependente funcional do campo MatFunc, ou seja, existe uma relação de dependência nos valores que preenchem estes campos.
- Sempre que o valor "F1" aparece no campo MatFunc, o valor "Helder" deve aparecer no campo NomeFunc.

Dependência Funcional

- Parcial: Quando a chave primária é composta, e existe um campo que depende somente de parte desta chave primária composta.
- Exemplo: a dependência existente entre NomeFunc e MatFunc.
 - NomeFunc é parcialmente dependente de MatFunc, que por sua vez faz parte da chave primária da tabela Funcionario_Projeto.

Dependência Funcional

- Transitiva: Quando uma coluna depende não somente da chave primária da tabela, mas também de uma segunda coluna da tabela.
- Exemplo: na tabela Funcionario_Departamento, o campo NomeDepto depende funcionalmente do campo NumDepto, que por sua vez depende da chave primária MatFunc.

2FN aplicado ao exemplo

Funcionario_Projeto

Matfunc	Numproj	Horas	NomeFunc	NomeProj
F1	PR3	20	Helder	Serra
F2	PR1	10	Antônio	Baixada
F3	PR2	30	Bruno	Z. Oeste
F1	PR2	15	Helder	Z. Oeste
F4	PR1	25	Amorim	Baixada
F3	PR3	10	Bruno	Serra

- MatFunc -> NomeFunc
- NumProj -> NomeProj

Funcionario_Projeto(Matfunc, NumProj, horas, NomeFunc, NomeProj)

Somente o campo **horas** depende da chave primária completamente

2FN aplicado ao exemplo

- Para resolver a dependência duas novas tabelas são criadas: Funcionário e Projeto.

Funcionario (Matfunc, NomeFunc)

Projeto (Numproj, NomeProj)

Funcionario_Projeto (Matfunc, Numproj, Horas)

Terceira Forma Normal (3FN)

- Tem o objetivo de eliminar as dependências funcionais transitivas e com isso eliminar possíveis redundâncias que ainda existam.
- Diz-se que uma entidade está na 3FN quando todos os seus atributos não chave não dependem de nenhum atributo não chave.

3FN aplicada ao exemplo

MatFunc	NomeFunc	Dtnasc	EndFunc	Numdepto	NomeDepto	MatGer
F1	Helder	31/08/76	R. do Rio, 72	D1	Logística	F5
F2	Antônio	18/12/85	R. Do Retiro, 10	D2	Qualidade	F6
F3	Bruno	03/07/79	R. Vermelha, 16	D1	Logística	F5
F4	Amorim	17/01/85	R. Monteiro, 25	D2	Qualidade	F6
F5	Artur	23/05/79	R. Paraná, 255	D2	Qualidade	F6
F6	Francisco	28/02/89	R. Acre, 12	D1	Logística	F5

Funcionario_Departamento

3FN aplicada ao exemplo

- Está na 1FN porque não possui dependência multivalorada;
- Está na 2FN porque não possui dependência funcional parcial, pois não possui chave primária composta;
- Entretanto há dependência entre os campos MatFunc e NumDepto, NumDepto e NomeDepto, e NumDepto e MatGer.

3FN aplicada ao exemplo

- Toda vez que o valor "F2" ocorre no campo MatFunc, o valor "D2" ocorre em NumDepto e toda vez que "D2" ocorre em NumDepto, ocorrerão valores "Qualidade" e "F6" para os campos NomeDepto e MatGer, respectivamente.
- O campo NumDepto não faz parte da chave primária da tabela.

Resultado da normalização

Funcionario (MatFunc, NomeFunc, DtNasc, EndFunc, NumDepto)

Departamento (NumDepto, NomeDepto, MatGer)

Funcionario_Projeto (MatFunc, Num_Proj, Horas)

Projeto (NumProj, NomeProj)

Quarta Forma Normal (4FN)

- Apesar do modelo estar na 3FN pode ocorrer alguma redundância. Isso ocorrerá quando um atributo não chave contiver valores múltiplos para uma mesma chave.

Quarta Forma Normal (4FN)

- Exemplo:

Musica	Interprete	Gravadora
Será	Renato Russo	EMI
Será	Simone	Polygram
Será	Renato Russo	Polygram
Imagine	John Lenon	EMI
Imagine	Simone	EMI
Imagine	John Lenon	Polygram

Note que não podemos criar uma chave com música nem com música e intérprete e, caso coloquemos os três campos como chave, teremos uma alternativa válida, mas redundante, pois haverá repetição de música e intérprete ou música e gravadora.

Quarta Forma Normal (4FN)

- Solução:
 - É necessária a criação de duas entidades: Gravadora e Intérprete. Ambos contendo o atributo música.

Quinta Forma Normal (5FN)

- Caso raro de ocorrer.
- Tecnicamente é utilizada a 5FN quando uma tabela na 4FN pode ser subdividida em duas ou mais tabelas, para evitar eventuais redundâncias.

Observação

- Este processo deve ser repetido tantas vezes quantas forem necessárias, até que não haja nenhum outro atributo ou grupo de atributos que estejam repetidos.