

BANCO DE DADOS I

Prof. Carlos Danilo Luz

FORMAS NORMAIS

O que seria isso de normalização?

Vamos imaginar que você se depara com a seguinte situação:

“Para realizar a estruturação das tabelas do Banco de Dados única coisa que temos é o relatório impresso do sistema antigo, agora precisamos recriar o mesmo”

A ideia de submeter um documento a normalização é para obter um modelo relacional a partir de qualquer tipo de representação de dados.

Com este processo, podemos identificar as tabelas e campos que poderão compor a estrutura de um banco de dados.

Estruturar em uma tabela não-normalizada os dados para obter uma descrição independente do tipo de documento que está sendo analisado. Tendo uma tabela base aplicamos uma sequência de passos que é denominada Normalização.

Para Heuser (2009), o objetivo da normalização é:

Reagrupar informações de forma a eliminar redundâncias de dados que possam existir nos arquivos.

Reagrupar informações de uma forma que permita a obtenção de um modelo ER.

Uma vez normalizados todos os arquivos do sistema, os diferentes esquemas relacionais obtidos pela normalização são integrados, gerando o esquema relacional do banco de dados do sistema (HEUSER, 2009).

A normalização utiliza-se de cinco etapas chamadas de formas normais (1FN, 2FN, 3FN, 4FN, 5FN), as três primeiras são as mais usuais.

RELATÓRIO DE ALOCAÇÃO A PROJETO

CÓDIGO DO PROJETO: LSC001

TIPO: Novo Desenv.

DESCRIÇÃO: Sistema de Estoque

<i>CÓDIGO DO EMPREGADO</i>	<i>NOME</i>	<i>CATEGORIA FUNCIONAL</i>	<i>SALÁRIO</i>	<i>DATA DE INÍCIO NO PROJETO</i>	<i>TEMPO ALOCADO AO PROJETO</i>
2146	João	A1	4	1/11/91	24
3145	Sílvia	A2	4	2/10/91	24
6126	José	B1	9	3/10/92	18
1214	Carlos	A2	4	4/10/92	18
8191	Mário	A1	4	1/11/92	12

CÓDIGO DO PROJETO: PAG02

TIPO: Manutenção

DESCRIÇÃO: Sistema de RH

<i>CÓDIGO DO EMPREGADO</i>	<i>NOME</i>	<i>CATEGORIA FUNCIONAL</i>	<i>SALÁRIO</i>	<i>DATA DE INÍCIO NO PROJETO</i>	<i>TEMPO ALOCADO AO PROJETO</i>
8191	Mário	A1	4	1/05/93	12
4112	João	A2	4	4/01/91	24
6126	José	B1	9	1/11/92	12

Após análise do documento fazemos o primeiro passo que podemos chamar de engenharia reversa, quando pegamos um documento e elaboramos uma suposta estrutura de dados, transformando documentos ou arquivos para passarem pela normalização.

O primeiro processo resulta em uma tabela que é dita não-normalizada. Esta que terá todos os dados aninhados, em sua visualização vamos perceber que aparecer valores de campos dentro de uma outra tabela, cada registro é como se estivesse em um vetor ou array de linguagens.

CódProj	Tipo	Descr	Emp					
			CodEmp	Nome	Cat	Sal	DataIni	TempAl
LSC001	Novo Desenv.	Sistema de Estoque	2146	João	A1	4	1/11/91	24
			3145	Sílvio	A2	4	2/10/91	24
			6126	José	B1	9	3/10/92	18
			1214	Carlos	A2	4	4/10/92	18
			8191	Mário	A1	4	1/11/92	12
PAG02	Manutenção	Sistema de RH	8191	Mário	A1	4	1/05/93	12
			4112	João	A2	4	4/01/91	24
			6126	José	B1	9	1/11/92	12

1 FORMAS NORMAIS

Com a tabela relacional não-normalizada correspondente do documento base, vamos iniciar o processo de normalização.

A passagem pela primeira forma normal (1FN) corresponde a transformação da tabela não-normalizada para um esquema relacional. Podemos considerar que uma tabela esteja na 1FN quando não contém tabelas aninhadas, ou seja, não contém dados estruturados como arrays.

Podemos transformar uma tabela não normalizada para a 1FN, tendo duas alternativas:

- Criação de uma tabela única, mas ciente que vai contar com redundância de dados.
- A segunda opção, é a mais interessante, é criar outra tabela para cada tabela aninhada. Assim os dados são separados em tabelas para melhorar a organização.

Na 1FN é realizado alguns passos básicos para a remoção dos dados aninhados, seguindo os seguintes passos:

- 1 - Criação de tabelas com base na tabela não-normalizada que contenham apenas colunas sem valores aninhados. A chave primária da tabela na 1FN seria basicamente a mesma utilizada na tabela não-normalizada.
- 2 - Ao separar as tabelas aninhadas, vamos ter basicamente a chave primária de cada uma das tabelas nas quais estava aninhada e as próprias colunas
- 3 - São identificadas as chaves primárias das tabelas que correspondem a cada tabela aninhada.

Proj:

CódProj	Tipo	Descr
LSC001	Novo Desenv.	Sistema
PAG02	Manutenção	Sistema de RH

ProjEmp:

CódProj	CodEmp	Nome	Cat	Sal	DataIni	TempAl
LSC001	2146	João	A1	4	1/11/91	24
LSC001	3145	Sílvio	A2	4	2/10/91	24
LSC001	6126	José	B1	9	3/10/92	18
LSC001	1214	Carlos	A2	4	4/10/92	18
LSC001	8191	Mário	A1	4	1/11/92	12
PAG02	8191	Mário	A1	4	1/05/93	12
PAG02	4112	João	A2	4	4/01/91	24
PAG02	6126	José	B1	9	1/11/92	12

2 FORMAS NORMAIS

A passagem pela segunda forma normal (2FN) tem como objetivo a eliminação de redundância de dados, como dito anteriormente, cada etapa da normalização é sequencial, ou seja, para que seja possível efetuar a 2FN da tabela deve já estar na 1FN.

Heuser (2009) afirma que:

Uma tabela encontra-se na segunda forma normal (2FN) quando, além de encontrar-se na primeira forma normal, cada coluna não chave depende da chave primária completa. Uma tabela que não se encontra na segunda forma normal contém dependências parciais, ou seja, contém colunas não chave que dependem apenas de uma parte da chave primária

Encontra-se na 2FN quando a tabela possuir uma chave primária simples, sendo também que na 2FN não se deve possuir dependências parciais, estas que ocorre quando uma coluna depende da composição da chave primária de outra coluna. vejamos a seguir a decomposição do processo de passagem para 2FN.

Proj:

CódProj	Tipo	Descr
LSC001	Novo Desenv.	Sistema de Estoque
PAG02	Manutenção	Sistema de RH

ProjEmp:

CódProj	CodEmp	DataIni	TempAl
LSC001	2146	1/11/91	24
LSC001	3145	2/10/91	24
LSC001	6126	3/10/92	18
LSC001	1214	4/10/92	18
LSC001	8191	1/11/92	12
PAG02	8191	1/05/93	12
PAG02	4112	4/01/91	24
PAG02	6126	1/11/92	12

Emp:

CodEmp	Nome	Cat	Sal
2146	João	A1	4
3145	Sílvio	A2	4
6126	José	B1	9
1214	Carlos	A2	4
8191	Mário	A1	4
4112	João	A2	4

3 FORMAS NORMAIS

A passagem pela 3FN ao final desta etapa já conseguimos extrair os dados para o modelo relacional, observe que a cada forma normal novas tabelas e estruturas são criadas, os dados são separados com o intuito de um melhor entendimento e cada vez evitar a redundância de dados.

Na 3FN a tabela passa por um processo para que não contenha dependências transitivas, uma dependência que ocorre quando uma coluna depende de uma chave primária e também de um conjunto de colunas da própria tabela.

Um tipo de redundância que se deseja eliminar no caso é mais complexa, pois depende da interpretação e da regra de negócio.

Exemplo: vamos imaginar que temos uma tabela de produtos, nesta temos a identificação do preço. Temos também uma tabela chamada categoria, a qual determina os preços dos produtos que estão na categoria X.

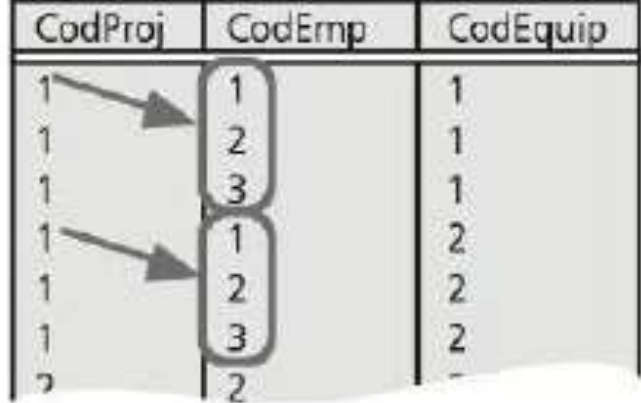
Perceba que tenho dois pontos de definição de preço, isso é o que desejamos remover. Para isso pode-se criar uma tabela associativa para armazenar apenas os dados de ligação.

4 FORMAS NORMAIS

A 4FN trabalha no conceito de dependência funcional multivalorada, onde temos uma redundância de dados muito específica.

Uma coluna ou conjunto de colunas depende multivaloradamente de uma coluna (determinante) da mesma tabela quando um valor do atributo determinante identifica repetidas vezes um conjunto de valores na coluna dependente (Heuser, 2009).

CodProj	CodEmp	CodEquip
1	1	1
1	2	1
1	3	1
1	1	2
1	2	2
1	3	2
2	2	2
2	2	4
3	3	1
3	4	1
3	3	3
3	4	3
3	3	5
3	4	5
4	2	5



CodProj	CodEmp	CodEquip
1	1	1
1	2	1
1	3	1
1	1	2
1	2	2
1	3	2
2	2	2

Podemos dizer que uma tabela está na 4FN quando fazemos a decomposição de uma tabela e esta não apresenta dependências multivaloradas, podendo assim criar novas tabelas para melhor compreensão dos dados.

ProjEquip:

CodProj	CodEquip
1	1
1	2
2	2
2	4
3	1
3	3
3	5
4	5

ProjEmp:

CodProj	CodEmp
1	1
1	2
1	3
2	2
3	3
3	4
4	2

A normalização é um processo complexo que se utiliza muito da abstração e interpretação dos dados, pois pensar na estruturação de tabelas partindo de um documento ou arquivo não é nada fácil. Geralmente criamos estas estruturas do zero, não aplicando uma engenharia reversa para suportar uma possível estrutura.

**VAMOS PRATICAR
FORMAS NORMAIS**

A normalização e utiliza-se de cinco etapas chamadas de formas normais (1FN, 2FN, 3FN, 4FN, 5FN), as três primeiras são as mais usuais.

Entidade: Filmes							
idFilme	Nome	Gênero	idMidia	Idioma	Tipo	Seção	Preço
656565	Uma janela suspeita	Drama	1001	dub	DVD	25	3,50
656565	Uma janela suspeita	Drama	1002	dub	VHS	25	3,50
656565	Uma janela suspeita	Drama	1003	leg	DVD	25	3,50
656565	Uma janela suspeita	Drama	1004	leg	VHS	25	3,50
323232	Minority Report	Ficção	2550	dub	DVD	32	4,20
323232	Minority Report	Ficção	2550	dub	VHS	32	4,20
323232	Minority Report	Ficção	2550	leg	DVD	32	4,20
323232	Minority Report	Ficção	2550	leg	VHS	32	4,20

Entidade: Filmes

idFilme	Nome	Gênero
656565	Uma janela suspeita	Drama
323232	Minority Report	Ficção

Entidade: Mídias

idMidia	idFilme	Idioma	IDTipo
1001	656565	151	15
1002	656565	151	17
1003	656565	171	15
1004	656565	171	17
2550	323232	181	18
2550	323232	181	19
2550	323232	191	18
2550	323232	191	19

Entidade: Tipo

IDTipo	Tipo	Preço
15	DVD	3,50
17	VHS	3,50
18	DVD	4,20
19	VHS	4,20

Entidade: SecaoIdioma

Idioma	Tipo	Seção
151	dub	25
171	leg	25
181	dub	32
191	leg	32

Banco de Dados I

Prof. Carlos Danilo Luz