



Aula 4. Normalização de dados

Objetivos:

- aprender a construir o Banco de Dados em sua estrutura; e
- produzir um projeto de Banco de Dados utilizando as técnicas de normalização.

Caro (a) estudante,

Nesta aula, aprenderemos a montar o Modelo Entidade Relacionamento (MER), considerando outro ponto de vista. Anteriormente, era apresentado o contexto em que tinha-se a ideia de modelar o banco e eram definidas as entidades, relacionamentos e atributos de forma bastante subjetiva. A normalização de dados trata da construção do modelo de dados, mas considera fatos reais, como documentos fiscais, formulários, planilhas, fichas, entre outras coisas. A ideia básica é que a partir destes documentos, aplicando regras previstas no processo de normalização, seja gerado o Modelo de Dados.

Para atingir os objetivos, podemos dizer que um banco de dados está relacionado a alguns conceitos fundamentais:

- Banco de Dados = Coleção de Arquivos;
- Arquivos = Coleção de Registros;
- Registro = Coleção de Campos (Tuplas);
- Campo = Coleção de Caracteres;
- Caractere = Alfa-numéricos ou símbolos.



De acordo com Machado (1995), existem nove técnicas que podem ser utilizadas no desenvolvimento de um banco de dados nas fases iniciais, as quais podemos citar:

- 1.** Contextualização;
- 2.** Objetivação;
- 3.** Intitulação;
- 4.** Especificação de Requisitos;
- 5.** Normalização;
- 6.** Modelagem;
- 7.** Trigramação;
- 8.** Dicionarização;
- 9.** Auditoria das fases de Análise e de Projeto Lógico do Sistema de Banco de Dados.

Para este curso, estaremos focados no que diz respeito à Normalização como ferramenta para o desenvolvimento de um banco de dados.

Os conceitos apresentados neste material, em seu todo ou em partes, foram retirados do livro “Projeto de Banco de Dados: uma visão prática”, de Machado, F. N. R. e Abreu, M. P., Editora Érica, 1995, Capítulo 12, Normalização.

4.1 Formas normais

A Teoria da Normalização é expressa tradicionalmente através de um conjunto de Formas Normais, que otimizam a estrutura e o conteúdo das relações entre as entidades.

O conceito de normalização foi introduzido por E. F. Codd, em 1970, utilizando a primeira forma normal – 1FN. Esta técnica, pode-se dizer, é um processo matemático formal que fundamenta-se na teoria dos conjuntos.



Nesse processo podem ocorrer as seguintes Anomalias de Atualização:

- Inclusão - ao se incluir um novo cliente, por exemplo, ele tem que estar relacionado com uma venda.
- Alteração - caso o fabricante de um produto altere o preço de determinada classe de produto, será preciso percorrer toda a Entidade para se realizarem múltiplas alterações.
- Exclusão - Ao se excluir um cliente, por exemplo, os dados referentes às suas compras poderão ser perdidos.

Através do processo de normalização podemos substituir, de forma gradual, um conjunto de Entidades e Relacionamentos para um modelo mais “adequado”, em relação às Anomalias de Atualização (Inclusão, Alteração e Exclusão), as quais podem causar certos problemas:

- Grupos repetitivos (atributos multivalorados) de dados;
- Dependências parciais em relação a uma chave concatenada;
- Redundâncias de dados desnecessários;
- Perdas acidentais de informação;
- Dificuldade na representação de fatos da realidade observada;
- Dependências transitivas entre atributos.

Estes problemas podem ser minimizados através da aplicação da normalização, tornando o modelo de dados bastante estável e sujeito a poucas manutenções.

4.2 Dependência funcional

Para descrever as formas normais, faz-se necessária a introdução deste conceito importante na aplicação da normalização, no qual a maior parte da teoria de normalização foi baseada.

Dada uma entidade qualquer, dizemos que o atributo ou conjunto de atribu-



tos de A é dependente funcionalmente de outro atributo de B; neste caso, há dependência funcional.

Em outras palavras, podemos dizer que, em uma entidade TB_ITENSNOTA, o atributo ITN_CODPRODUTO depende funcionalmente do atributo PRO_CODIGO, que encontra-se na entidade TB_PRODUTO.

A análise das relações existentes entre os atributos de uma entidade deve ser feita a partir de um conhecimento prévio da sistemática aplicada nas relações entre as entidades, para que esta possa ser modelada, como no exemplo anterior.

Para iniciar a normalização, faz-se necessária a descrição de todos os atributos do documento que se deseja normalizar (visando a criação do modelo de dados), para que posteriormente, possam ser aplicadas técnicas (formas normais) que irão dividir os atributos em entidades, as quais darão suporte à construção do modelo de dados.

Para facilitar o entendimento, este material irá apresentar a normalização com exemplos práticos.



4.3 Normalização de dados (nota fiscal)

Supondo a necessidade de uma empresa em modelar um banco de dados para emissão de uma nota fiscal simples, como apresentado no quadro abaixo. Verifique que há dados que são específicos da empresa, representados aqui como "Dados da Empresa" e que não há necessidade de manter-se no banco de dados, pois são impressos diretamente na nota fiscal encaminhada pela gráfica. O que nos interessa aqui é justamente os dados que necessitam ser preenchidos, como Data, Número da Nota Fiscal, Cliente, RG, CPF, Cidade, UF, Endereço, Número e Bairro, além dos dados dos produtos vendidos, inerentes ao Código do produto, Descrição, Unidade de Medida, Quantidade, Valor Unitário e Valor Total.



Dados da Empresa		Data: 01/01/0101		NF N° 123
CLIENTE: João da Silva		RG: 0000000000-00		
CPF:000000000-00		CIDADE: Medianeira		UF: PR
ENDEREÇO: R. Bahia		Nº 1234	BAIRRO: Centro	
Código	Descrição	Und. de Medida	Quantidade	Vlr Unitário
1	Arroz	kg	2	1,99
2	Sabonete	Unidade	1	0,79
3	Leite	Litro	3	2,12
Valor TOTAL : 11,13				

Inicialmente, realiza-se a definição de todos os atributos que o documento possui, registrando que eles são tidos como atributos da entidade principal (documento), conforme segue.

Neste caso, a entidade principal foi definida como Nota e a ela foram relacionados todos os atributos oriundos a emissão de uma nota fiscal. Foi definida ainda a chave primária, única para cada nota fiscal a ser emitida.

Nota (número da nota, data, cliente, RG, CPF, cidade, UF (Unidade da Federação), endereço, nº, bairro, código_produto_1, código_produto_2, código_produto_3, ..., descrição_produto_1, descrição_produto_2, descrição_produto_3,..., unidade_produto_1, unidade_produto_2, unidade_produto_3, quantidade_produto_1, quantidade_produto_2, quantidade_produto_3, ..., valor_unitário_produto_1, valor_unitário_produto_2, valor_unitário_produto_3, valor_total_produto_1, valor_total_produto_2, valor_total_produto_3,..., valor_total_da_nota);

- Os campos sublinhados indicam as chaves primárias.

4.3.1 Aplicação da 1ª Forma Normal – 1FN

Existem situações em que nos deparamos com algumas informações que se repetem (atributo multivalorado) dentro de uma única linha, ligada a uma chave primária.

A 1FN diz que cada ocorrência da chave primária deve corresponder a somente uma informação de cada atributo, ou seja, a entidade não deve conter atributos repetidos ou multivalorados, ou, ainda, os atributos não-chave deverão ser atônicos (unicos).



Ao observarmos que certo grupo de atributos não-chave não são atômicos, ao longo do processo de entrada de dados, verificamos que este atributo deverá ser decomposto em uma nova entidade.

Nas novas entidades criadas, a chave primária será a concatenação da chave primária da entidade original, o que mantém, desta forma, o relacionamento entre elas, como também o conjunto de atributos que se repetem.

Por exemplo:

A 1FN diz que todos os atributos devem conter um valor atômico (único), ou seja, os dados inclusos dentro do campo não devem se repetir.

Ex.: o campo cliente vai ser preenchido quantas vezes na nota acima? Apenas uma vez, certo? Então corresponde a um valor atômico. Já os campos que se referem aos itens da nota serão acrescidos uma ou muitas vezes. Desta forma, tem-se a necessidade de separação da entidade Nota, criando-se a nova entidade chamada Itens_Nota, ficando:

Nota (número da nota, data, cliente, RG, CPF, endereço, nº, bairro, cidade, UF, valor_total)

Os campos que se repetem são destinados a uma nova tabela, no nosso exemplo chamado de itens da nota. A chave principal também é inclusa na nova tabela para manter o relacionamento entre elas.

Itens_Nota (número da nota, código, descrição, unidade_de_medida, valor_unitário, quantidade, valor_total)

- Note: as tabelas que saem na 1FN são tabelas associativas.

4.3.2 Aplicação da 2^a Forma Normal – **2FN**

A 2FN diz que todos os atributos que não forem chave têm que ser dependentes diretos da chave, ou seja, os atributos que estão contidos na nota têm que depender diretamente da chave **Número da nota**.

Inicialmente, observa-se a entidade que possui chave primária concatenada. Para aquelas que satisfazem esta condição, analisar se existe algum atributo ou conjunto de atributos com dependência direta à chave primária.



Ex.: o atributo cliente é dependente direto da nota, pois é o cliente desta nota. Já atributo RG, CPF são dependentes do cliente, não tendo nenhuma relação direta com a nota, ou seja, independentemente da nota, o CPF do cliente será o mesmo se soubermos o nome do cliente. Sabe-se também que o CPF do cliente nunca muda. Desta forma, tem-se:

Nota (número_da_nota, data, código_cliente, endereço, nº, bairro, código_cidade, valor_total);

- Os atributos adicionados acima “código_cliente e código_cidade” são para manter o relacionamento entre as entidades.
- Verifique que o campo endereço deve permanecer na entidade Nota, tendo em vista que, se, por acaso, pretender-se buscar uma nota emitida há cinco anos cujo cliente alterou seu endereço para outra rua, o endereço que deverá aparecer na nota emitida há cinco anos deve ser o correspondente ao endereço antigo do cliente, e não o novo endereço.

Cliente (código, nome, RG, CPF);

- Os campos dependentes do cliente são transferidos para a tabela cliente.

Cidade (código, descrição, UF);

- O atributo UF que depende apenas da cidade é atribuído para a tabela cidade.

Itens nota (número_da_nota, código_produto, quantidade, valor_unitário, valor_total);

- O atributo “código_produto” é mantido na tabela associativa para manter o relacionamento entre Nota, Itens da Nota e Produtos.

Produto (código, descrição, unidade_de_medida, valor_unitário);

- Os atributos dependentes de produto são relacionados tabela produto. Repare que o atributo valor unitário foi duplicado, isso pela necessidade do valor unitário de um produto sofrer mudanças, em notas diferentes o valor unitário pode estar diferente também.



4.3.3 Aplicação da 3ª Forma Normal - 3FN

A 3FN diz que o atributo precisa estar na segunda forma normal, e todos os atributos que não são chave, não podem depender de outros atributos que também não são chave. Deve-se verificar ainda se o valor de um atributo pode se repetir em muitas tuplas, sendo que neste caso, é necessário inclusão de nova entidade.

Uma Entidade está na 3FN se nenhum de seus atributos possui dependência em relação a outros atributos da Entidade. Ao verificar a existência de dependência entre as Entidades e seus atributos, devemos criar outra entidade que contenha os atributos dependentes.

Nota (número da nota, data, código_cliente, código_endereço, código_bairro, nº, código_cidade);

- Note que o atributo valor total da nota foi retirado na 3FN. Todos os atributos que são gerados a partir do resultado de dois ou mais atributos são retirados, visando manter a consistência dos dados.

Endereço (código, descrição);

- Aplicando a 3FN, o endereço é retirado da tabela nota, pois o endereço não é um atributo chave e refere-se ao endereço do cliente nesta nota. Então, o endereço é incluso na tabela de clientes e relaciona-se tanto com a entidade Cliente como a entidade Nota.

Bairro (código, descrição);

- Aplicando a 3FN, o bairro é retirado da tabela nota, pois o bairro não é um atributo chave e refere-se ao bairro do cliente nesta nota. Então, o bairro é incluso também na tabela de clientes e relaciona-se tanto com a entidade Cliente como a entidade Nota.

Cliente (código, nome, RG, CPF, nº, código_endereço, código_cidade);

- Os campos chave código_endereço e código_cidade vêm para a tabela de clientes com o fim de fazer o relacionamento entre as tabelas.



Cidade (código, descrição, UF);

Itens nota (número da nota, código produto, quantidade, valor_unitário);

- O atributo valor total sai, pois ele é resultado de outros dois atributos: a quantidade vezes o valor unitário.

Produto (código, descrição, unidade_de_medida, valor_unitário);

A figura 18 representa o MER do exemplo normalizado anteriormente.

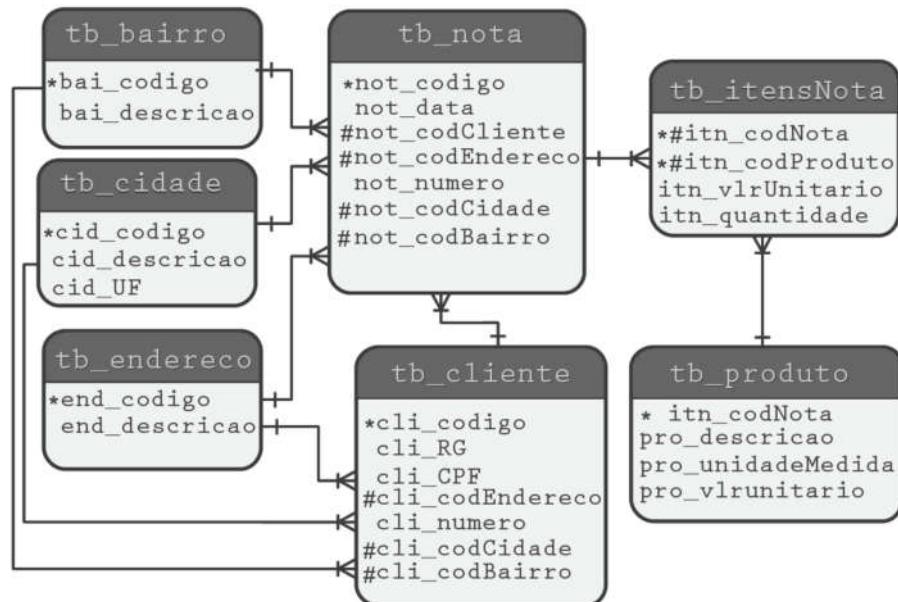


Figura 18 - Modelo gerado pela Normalização da Nota Fiscal

Fonte: autor (adaptado pelo ilustrador)



4.4 Normalização de dados (Formulário Remanejamento)

Dado o formulário de solicitação de remanejamento, será aplicada a normalização de dados visando obter um modelo de dados conciso:

	MINISTÉRIO DA EDUCAÇÃO UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ CAMPUS MEDIANEIRA
REQUERIMENTO DE REMANEJAMENTO	
ESPAÇO A SER PREENCHIDO PELO REQUISITANTE	
Nome:	
Sexo: <input type="checkbox"/> M <input type="checkbox"/> F Data de Nascimento:/...../.....	
Categoria Funcional:	
Lotação atual: Data de admissão:/...../..... Cidade:	
Escolaridade <input checked="" type="checkbox"/> 1º Grau <input type="checkbox"/> 2º Grau - Curso	
<input type="checkbox"/> 3º Grau - Curso:	
<input type="checkbox"/> Outros:	
Principais cursos realizados: 1..... 2..... 3..... 4..... 5.....	
Experiências profissionais relevantes:	
Principais atribuições desenvolvidas ao setor:	

Figura 19

Fonte: Machado (1995)



Setores em que já trabalhou:

1Período:a

Motivo da saída:.....

2Período:a

Motivo da saída:.....

3Período:a

Motivo da saída:.....

Obs: se necessário, utilizar folha avulsa e anexar

Atribuições que gostaria de trabalhar, ordem de preferência:

1.....

2.....

3.....

4.....

5.....

Horário de preferência:

Setores/Divisões/ Diretorias em que gostaria de trabalhar, em ordem de preferência:

1.....

2.....

3.....

Qual(is) o(s) motivo(s) que o torna descontente com suas atividades?

- Falta de correlação com a formação acadêmica
- Atividade com complexidade mínima
- A falta de formação em relação à atividade desenvolvida prejudica sua execução
- Horário de trabalho
- O relacionamento com a chefia não é satisfatório
- O relacionamento com os colegas de trabalho não é bom
- Faz a mesma coisa há muitos anos
- Falta de expectativa de crescimento profissional
- A atividade desenvolvida está criando problemas de saúde
- Excesso de trabalho
- Poucos trabalho
- Ambiente excessivamente autocrático

Outros.....

Observações:

.....
.....
.....

Data de preenchimento:...../...../.....

Medianeira, de de

.....

Assinatura do servidor

Figura 20
Fonte: Machado (1995)



ESPAÇO A SER PREENCHIDO PELA CHEFIA IMEDIATA

O(a) servidor(a) tem demonstrado interesse pela atividade atualmente desenvolvida?

.....

As tarefas que lhe são confiadas são cumpridas a contento?

.....

Os horários cumpridos estão dentro do estabelecido pela chefia?

.....

O(a) servidor(a) respeita os regulamentos e age em conformidade com as normas internas?

.....

O(a) servidor(a) desempenha suas funções com seriedade e precisão?

.....

Qual a posição da chefia em relação à permanência do servidor no atual setor?

.....

Qual o relacionamento do(a) servidor(a) com o grupo de trabalho?

.....

Mediante as informações acima prestadas:

- () Libero o(a) funcionário(a) sem reposição
() Libero o(a) funcionário(a) mediante reposição
() Não libero o(a) funcionário(a)

Justificativa:

.....

.....

Data:/...../..... Assinatura e carimbo:

Espaço a ser preenchido pela Diretoria em que o servidor está lotado

Data:/...../..... Assinatura e carimbo:

Espaço a ser preenchido pela Divisão de Recrutamento, Seleção e Desenvolvimento de Pessoal

Figura 21

Fonte: Machado (1995)



4.4.1 Definição dos atributos contidos no formulário

Requerimento (código, nome, sexo, data_nascimento, categoria_funcional, lotação_atual, data_admissão, cidade, escolaridade, grau, cursos_realizados, experiências_profissionais, atribuições_desenvolvidas, setor_que_trabalhou, data_entrada, data_saída, motivo_saída, atribuições_que_gostaria, horário, setores/divisões/diretorias_gostaria, motivos_descontente, observações, data_preenchimento, questão1, questão2, questão3, questão4, questão5, questão6, questão7, resultado, justificativa, data_resultado)

- Os campos sublinhados indicam as chaves primárias.

4.4.2 Aplicação da 1ª Forma Normal - 1FN

A 1FN diz que todos os atributos devem conter um valor atômico (único), ou seja, os dados inclusos dentro do campo não devem se repetir.

Ex.: o campo que representa os cursos feitos pelo requerente será preenchido quantas vezes no formulário? Depende de quantos cursos ele fez: na possibilidade de ser preenchido duas vezes, o campo é retirado para uma tabela associativa.



Remanejamento (código, nome, sexo, data_nascimento, categoria_funcional, lotação_atual, data_admissão, cidade, data_preenchimento, questão1, questão2, questão3, questão4, questão5, questão6, questão7, resultado, justificativa, data_resultado);

- Permaneceram na tabela remanejamento apenas os campos com valores únicos, os campos que podem ser preenchidos uma ou mais vezes são retirados para as tabelas associativas.

Remanejamento/escolaridade (código_remanejamento, escolaridade, grau);

Remanejamento/Cursos (código_remanejamento, cursos);

Remanejamento/experiências profissionais (código_remanejamento, experiências_profissionais);

Remanejamento/atribuições desenvolvidas (código_remanejamento, atribuições);



Remanejamento/setores que trabalhou (código_remanejamento, setor, data_entrada, data_saída, motivo_saída);

Remanejamento/atribuições que gostaria (código_remanejamento, atribuições, horário);

Remanejamento/setores-divisões-diretorias que gostaria (código_remanejamento, setor);

Remanejamento/motivos descontentamento (código_remanejamento, motivos, observações).

4.4.3 Aplicação da 2ª Forma Normal - 2FN

A 2FN diz que todos os atributos que não forem chave têm que ser dependentes diretos da chave, ou seja, os atributos que estão contidos no requerimento têm que depender diretamente da chave código requerimento.

Inicialmente, observa-se a entidade que possui chave primária concatenada. Para aquelas que satisfazem esta condição, analisar se existe algum atributo ou conjunto de atributos com dependência direta à chave primária.

Remanejamento (código, código_funcionário, categoria_funcional, lotação_atual, data_admissão, cidade, data_preenchimento, questão1, questão2, questão3, questão4, questão5, questão6, questão7, resultado, justificativa, data_resultado);

- Verifique que o campo categoria funcional, lotação atual e cidade devem permanecer na entidade Remanejamento, tendo em vista que, se, por acaso, se desejar buscar um formulário emitido há cinco anos, cujo funcionário alterou sua categoria funcional, lotação atual ou cidade, os dados que deverão aparecer no formulário preenchido há cinco anos devem ser o correspondente à categoria funcional, lotação atual ou cidade, preenchido antigamente, e não os dados novos.

Funcionário (código, nome, sexo, data_nascimento);

Remanejamento/escolaridade (código_remanejamento, escolaridade, grau);

Remanejamento/Cursos (código_remanejamento, cursos);



Remanejamento/experiências profissionais (código_remanejamento, experiências);

Remanejamento/atribuições desenvolvidas (código_remanejamento, atribuições);

Remanejamento/setores que trabalhou (código_remanejamento, setor, data_entrada, data_saída, motivo_saída);

Remanejamento/atribuições que gostaria (código_remanejamento, atribuições, horário);

Remanejamento/setores-divisões-diretorias que gostaria (código_remanejamento, setor);

Remanejamento/motivos descontentamento (código_remanejamento, motivos, observações);

4.4.4 Aplicação da 3ª Forma Normal – 3FN

A 3FN diz que o atributo precisa estar na segunda forma normal, e todos os atributos que não são chave, não podem depender de outros atributos que também não são chave.

Uma Entidade está na 3FN se nenhum de seus atributos possui dependência em relação a outros atributos da entidade. Ao verificar a existência de dependência entre as Entidades e seus atributos, devemos criar outra entidade que contenha os atributos dependentes.

Remanejamento (código, código_funcionário, código_categoria_funcional, código_lotação, data_admissão, código_cidade, data_preenchimento, questão1, questão2, questão3, questão4, questão5, questão6, questão7, resultado, justificativa, data_resultado);

Categoria Funcional (código, descrição);

Lotação (código, descrição);

Cidade (código, descrição);

Funcionário (código, nome, sexo, data_nascimento);



Remanejamento/escolaridade (código_remanejamento, código_escolaridade, grau);

Escolaridade (código, descrição);

Remanejamento/Cursos (código_remanejamento, código_cursos);

Cursos (código, descrição);

Remanejamento/experiências profissionais (código_remanejamento, código_experiências);

Experiências Profissionais (código, descrição);

Remanejamento/atribuições desenvolvidas (código_remanejamento, código_atribuições);

Atribuições (código, descrição);

Remanejamento/setores que trabalhou (código_remanejamento, código_setor, data_entrada, data_saída, motivo_saída);

Setores (código, descrição);

Remanejamento/atribuições que gostaria (código_remanejamento, código_atribuições, horário);

Remanejamento/setores-divisões-diretorias que gostaria (código_remanejamento, código_setor);

Remanejamento/motivos descontentamento (código_remanejamento, código_motivos, observações);

Motivo descontentamento (código, descrição).

A figura 22 demonstra o MER do exemplo anterior.

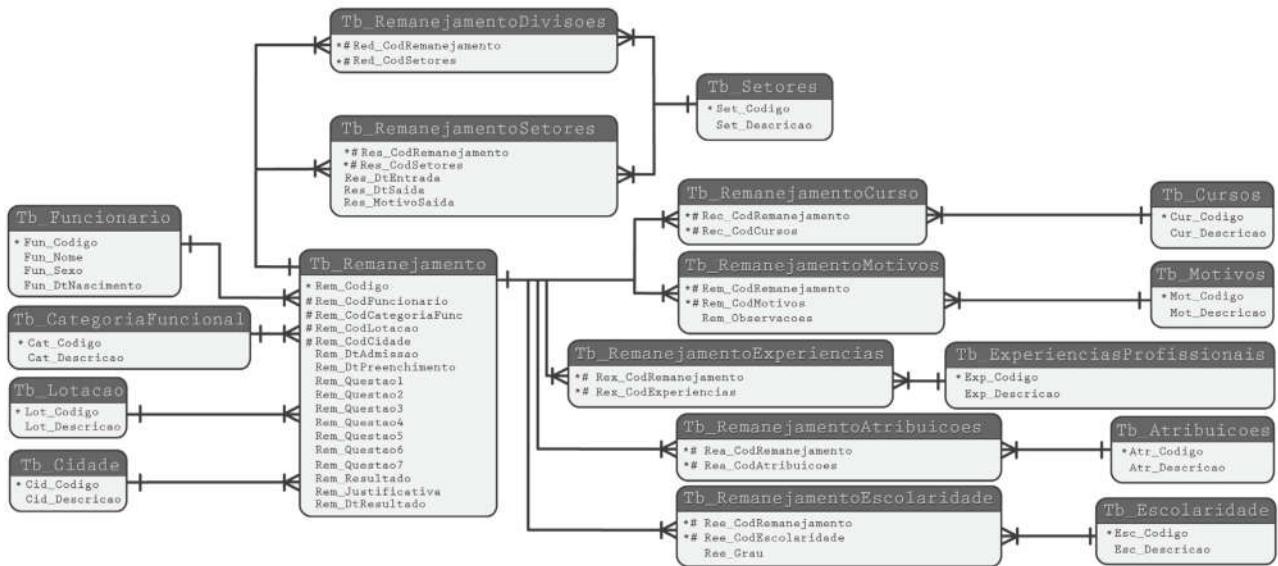


Figura 22 - Modelo de Dados aplicado ao Requerimento de Remanejamento

Fonte: autor (adaptado pelo ilustrador)

Resumo

Chegando ao fim deste capítulo, você deve ter percebido que a normalização de dados é uma ferramenta a mais, que propicia a criação de um Modelo de Dados conciso de forma facilitada. Lembre-se de que sempre há necessidade de seguir as regras "ao pé da letra", considerando que caso não faça isso, seu modelo gerado poderá apresentar problemas, como duplicidade de dados, inconsistência, entre outros. Fixe a ideia de que a primeira forma normal busca determinar todas as tabelas associativas que você terá. A segunda forma normal busca manter a integridade do banco, mantendo inclusive o aspecto histórico necessário e sua dependência com as entidades. A terceira forma visa criar entidades que normalmente não dependem de ninguém, mas que se tornam dependentes de outras entidades.