

4 Modelagem de Dados

4.1 A importância da modelagem de dados

Quando pretendemos fazer uma viagem, a primeira preocupação é planejá-la cuidadosamente e escolher o que vamos levar nas malas. Da mesma forma, quando estamos empenhados no desenvolvimento de um sistema para computador, é preciso planejar todas as suas etapas e dedicar atenção especial ao projeto e estruturação do banco de dados. Para isso utilizamos uma técnica chamada modelagem de dados, cujo objetivo é transformar uma idéia conceitual em algo que possa ser traduzido em termos computacionais. Com a modelagem de dados é possível refinar um modelo conceitual durante as fases que compõem o projeto, eliminando redundâncias ou incoerências que possam inevitavelmente surgir.

Sem esse planejamento prévio, com toda certeza, a manutenção do sistema será mais complicada e constante. Não podemos pensar que todo o projeto do banco de dados deva ser assumido somente pela equipe de profissionais de informática. Muito pelo contrario, é preciso que os próprios usuários façam parte das etapas mais críticas como levantamento dos dados, testes de usabilidade e validação. Não há dúvida de que os profissionais também precisam conhecer o negócio da empresa, ou poderão aparecer resultados desastrosos.

Uma vez que a realidade muda de uma empresa para outra, é preciso estabelecer uma forma-padrão para estruturar um banco de dados independente do tipo de ambiente ou negócio. Para isso foi definido o que chamamos de “metamodelo”, sendo que o mais utilizado é **Entidade-Relacionamento** (ER).

O processo de desenvolvimento de um projeto de banco de dados envolve varias fases. O mini mundo representa uma porção da realidade que se deseja “transportar” para o sistema informatizado. Na fase de levantamento e análise de necessidades a equipe de desenvolvimento deve consultar os potenciais usuários para saber quais são os requisitos básicos do banco de dados e assim documentá-los adequadamente. A partir dessa fase são gerados dois produtos, os requisitos funcionais que englobam todas as funções da aplicação, e os requisitos do banco de dados.

O próximo passo é a criação de um esquema conceitual do banco de dados valendo-se de um modelo de dados de alto nível. Esse esquema consiste numa descrição detalhada dos requisitos de dados apontados no levantamento junto aos usuários. Nessa descrição encontram-se detalhes sobre as entidades, relacionamentos e restrições.

Na implementação propriamente dita se encontram as tarefas de criação do banco de dados físico, utilizando um dos produtos oferecidos no mercado, e na programação das rotinas e transações que o aplicativo deve conter a partir de ferramentas de desenvolvimento e linguagens disponíveis.

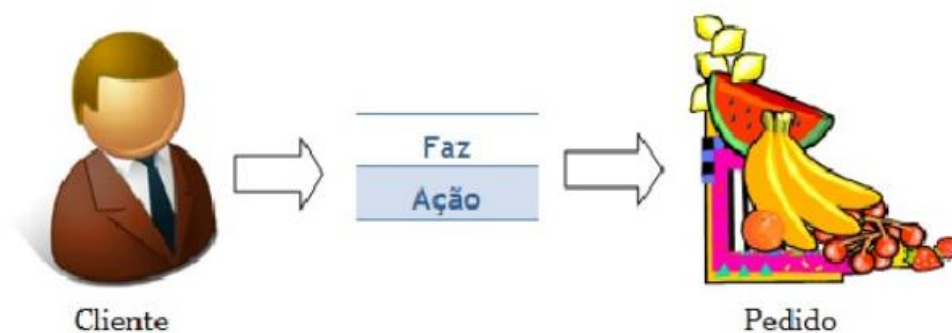
É no projeto físico que são especificados a estrutura do banco de dados, os caminhos para acesso e a forma de organização dos arquivos que o compõem.

No processo de modelagem de dados podemos destacar como básicas as seguintes etapas:

1. Definição do problema ou objetivo principal;
2. Pesquisa dos sistemas de banco de dados existentes no mercado para adoção;
3. Projeto da estrutura de dados;
4. Construção dos relacionamentos entre os dados;
5. Implementação de regras de negócio e restrições;
6. Implementação do desenho da interface com usuário, que também envolve a criação de relatórios do sistema.

4.2 Conceito de Entidade-Relacionamento ER

Esse modelo foi concebido em 1976, por Peter Chen, com base na teoria de bancos de dados relacionais de Edgard F. Codd. O conceito principal por trás do modelo E-R (Entidade-Relacionamento) estão na definição de dois grupos de objetos que formam um negócio: Entidades e Relacionamentos. Eles possuem uma ligação tão forte que não é possível tratar de um sem mencionar o outro. O que une esses dois componentes é uma ação. É como uma oração em que temos um sujeito (entidade), um verbo (ação) e um complemento (relacionamento).



Durante a fase inicial de modelagem conceitual dos dados, o profissional precisa observar atentamente tudo que for relevante no mundo real e que deva ser “transportado” para o sistema que se está projetando. Com essas informações já é possível criar um esboço representando de forma gráfica o processo. A isso denominamos **abstração** ou **modelo abstrato**. Nela podemos encontrar três componentes muito importantes: modelo conceitual, modelo lógico e modelo físico.

4.2.1 Modelo Conceitual

É a primeira etapa do projeto, na qual se representa a realidade através de uma visão global e genérica dos dados e seus relacionamentos. Seu objetivo é conter todas as informações dessa realidade que serão armazenadas no banco de dados, sem que se retratem aspectos relativos ao banco de dados que será utilizado. Essas informações podem aparecer no formato de uma lista descritiva das operações executadas pelos usuários e os dados que eles devem manipular. Veja a seguir um pequeno exemplo:

1. Cadastro de Clientes

Dados necessários: nome completo, tipo de pessoa (física/jurídica), endereço, bairro, cidade, estado, CEP, telefone, e-mail, nome para contato.

2. Cadastro de fornecedores

Dados necessários: nome completo, tipo de pessoa (física/jurídica), endereço, bairro, cidade, estado, CEP, telefone, e-mail, nome para contato.

3. Cadastro de produtos

Dados necessários: código do produto, nome do produto, modelo, unidade de medida, preço unitário de compra, preço unitário de venda, estoque mínimo, estoque Máximo, estoque atual, data última compra.

4. Pedidos

Dados necessários: código do produto, quantidade, código do cliente, código do vendedor.

5. Operações

a. Cadastro de produtos

O funcionário adiciona um registro quando um novo produto é comprado, incluindo sua descrição, preço de compra e de venda e a quantidade que está entrando em estoque.

b. Cadastro de clientes

O funcionário adiciona um registro quando o cliente que está efetuando um pedido a prazo não possui cadastro.

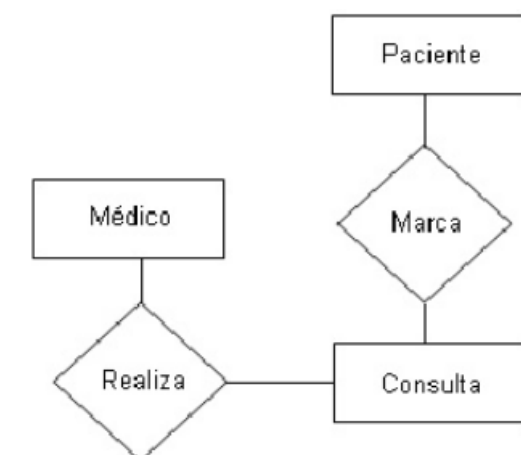
c. Vendas

O funcionário entra com os produtos que o cliente deseja comprar. Para finalizar a venda, deve-se especificar se é a vista ou a prazo. Na finalização da venda, a nota fiscal é impressa e o estoque dos produtos é atualizado. Se for venda a prazo, um carnê é emitido para pagamento das parcelas mensais.

4.2.2 Modelo Lógico

A segunda etapa compreende a descrição das estruturas que serão armazenadas no banco de dados e que resulta numa representação gráfica dos dados numa maneira lógica, inclusive já nomeando os componentes e as ações que exercem um sobre o outro.

Nessa etapa também se define a abordagem de banco de dados que será utilizada: hierárquica, de rede ou relacional.



4.2.3 Modelo Físico

Do **modelo lógico** podemos derivar o **modelo físico**, no qual se encontram detalhados os componentes de estrutura física do banco de dados, como tabelas, campos, tipos de valores, índices, etc.

Quando chegarmos a esse ponto, estaremos prontos para a criação propriamente dita do banco de dados, utilizando o sistema gerenciador que mais se adequar as tarefas. Veja um exemplo hipotético:

Estrutura do Cadastro de Paciente

Nome do campo	Tipo de dado	Tamanho do Campo
Código do Paciente	Númérico	5 dígitos
Nome do Paciente	Alfanumérico	50 caracteres
Endereço	Alfanumérico	50 caracteres
Bairro	Alfanumérico	40 caracteres
Cidade	Alfanumérico	40 caracteres
Estado	Alfanumérico	2 caracteres
CEP	Alfanumérico	9 caracteres
Data de Nascimento	Data	10 caracteres

4.3 Entidades

No modelo ER (Entidade-Relacionamento) os dados são descritos como entidades, atributos e relacionamentos, que podem ser exibidos num formato gráfico denominado diagrama ER.

Sendo a entidade o objeto básico representado no modelo ER, veremos seu conceito em primeiro lugar. Podemos definir entidade como um objeto do mundo real que possui atributos capazes de torná-lo identificável e que possui uma existência independente. Essa existência pode ser física— como no caso de pessoas, casa, relógio, computadores, funcionários, etc.— ou apenas conceitual— como um serviço, uma disciplina escolar, uma consulta medica, etc.

As entidades possuem atributos que as descrevem, como a entidade de um produto, que é representada pelo seu código, nome, modelo, preço e estoque. Esse atributo são partes importantes dos valores armazenados no banco de dados.

Os atributos podem ser classificados em simples ou compostos. Quando a informação contida no atributo não pode ser dividida, como o preço unitário de um produto, então ele é tido como simples. Já os atributos compostos permitem uma divisão em partes menores da informação. Tomemos como exemplo a entidade de endereços, que pode ser quebrada em nome do logradouro, número do imóvel, bairro, cidade, estado e CEP.

Endereços					
Logradouro	Número imóvel	Bairro	Cidade	Estado	CEP

Mas como podemos reconhecer uma entidade?

Isso pode ser feito pela análise criteriosa de cada informação levantada. É nessa análise que se procura agrupar as informações conforme suas características, ou que estão relacionadas com um mesmo assunto.

Vamos supor o caso de um cadastro de funcionários de uma empresa. Além da lista de funcionários, temos também uma relação de setores em que cada um trabalha. Dessa forma podemos identificar de imediato dois grupos de informações.



Como foram já mencionadas, as entidades possuem atributos que diferenciam uma das outras. Esses atributos contem valores que descrevem as diversas ocorrências de uma entidade, que são identificadas como os registros de uma tabela no banco de dados físico.

Entidade Setor

Código do Setor	Nome do Setor
0001	Administração
0002	Tesouraria
0003	Atendimento
0004	Almoxarifado
0005	Contabilidade

Entidade Funcionário

Matrícula	Nome do Funcionário	Código de Setor
20145-8	José Bento Figueira	0002
13002-1	Augustino Valentin	0003
09021-4	Alberto Vasconcelos	0004
80240-1	Denise Souza	0003
84230-8	Bruno Lima	0004
12090-3	Iris da Costa	0005

Existe uma característica dos atributos que indica quantas vezes ele pode ocorrer, a qual recebe o nome de cardinalidade. A maioria dos atributos somente armazena um único valor, e chamamos isso de cardinalidade 1. Este é o caso do atributo do nome do funcionário. Em algumas situações, no entanto, é possível termos atributos que devem conter mais de um valor, que denominados de atributos multivalorados. O atributo para armazenamento da formação profissional do funcionário é um exemplo, já que pode conter um ou mais valores, no caso de se ter vários cursos superiores.

Os atributos também são formados por um tipo de dado (numérico, caractere, data, etc.) e por um tamanho que representa a quantidade de dígitos ou caracteres alfanuméricos que podem ser armazenados.

Um atributo especial (ou conjunto de atributos) pode representar uma ocorrência única de entidade dentro de toda a coleção. É a chave da entidade, que no banco de dados físico é representada pela chave primária.

As entidades que não possuem atributos chaves são denominadas de **tipos de entidades fracas**, ao contrario dos **tipos de entidades fortes**, que são aquelas que possuem atributos chaves. Os itens/registros de entidades que pertencem ao primeiro tipo se relacionam com itens/registros existentes em outra entidade, sendo essa ultima denominada **entidade identificadora**.

4.4 Relacionamentos

Um relacionamento pode ser definido como a representação de uma ação ou fato que associa os itens de uma entidade com os itens de outra entidade. Falando em termos matemáticos, um relacionamento pode ser parecer com as linhas que ligam os elementos de um conjunto aos elementos de outro. Na modelagem de dados, os relacionamentos são representados graficamente como um losango.

Vamos utilizar o exemplo anterior das entidades **SETOR** e **FUNCIONÁRIO**. Nele podemos perceber que para ser funcionário da empresa ele deve estar “lotado” num determinado setor. Assim temos um vínculo entre essas duas entidades.



Como podemos notar, há um verbo (de preferência na voz ativa) ligando as duas entidades que descrevem com exatidão como elas se relacionam. As ocorrências da entidade **SETOR** se associam com as ocorrências de entidade **FUNCIONÁRIO**. Da mesma forma as ocorrências de **FUNCIONÁRIO** “estão lotadas” nas ocorrências de **SETOR**.

4.4.1 Condicionalidade

A Condicionalidade refere-se à capacidade de uma entidade poder ter ou não ligação/vínculo com outra. Isso significa que pode haver ocorrências numa entidade que não possuem um relacionamento ou associação com ocorrências de outra entidade. No exemplo de cadastro de funcionário, há uma obrigatoriedade de vincular um funcionário a um setor. No entanto, não é obrigatório que todos eles tenham filhos, o que indica que o relacionamento da entidade **FUNCIONÁRIO** com outra entidade denominada **FILHO** nem sempre ocorrerá.

Só porque alguns itens não participam do relacionamento não significa que o fato não possa existir. Isso define dois grupos de relacionamentos: condicionais e incondicionais.

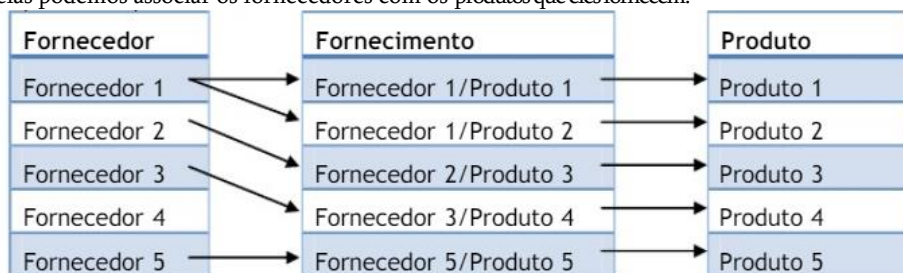
Nos relacionamentos condicionais pode haver elementos que não participam do relacionamento, ao contrario dos relacionamentos incondicionais, em que o relacionamento dos elementos entre as entidades é obrigatório, ou seja, todos os elementos de uma entidade devem ser associar com pelo menos um elemento de outra entidade.



4.4.2 Tipos e Graus de Relacionamento

Quando num relacionamento temos envolvidas duas entidades, dizemos que ele é binário. Se forem três entidades, temos um relacionamento ternário. Até agora todos os exemplos eram do primeiro tipo.

Como exemplo de relacionamento ternário pode citar o caso em que temos três entidades: **Fornecedor**, **Produto** e **Fornecimento**. Por meio delas podemos associar os fornecedores com os produtos que eles fornecem.



O **grau de relacionamento** indica a quantidade de ocorrências de uma entidade que se encontram relacionadas com ocorrências de outra entidade. Há três tipos de grau de relacionamento: um-para-um; um-para-muitos e muitos-para-muitos.

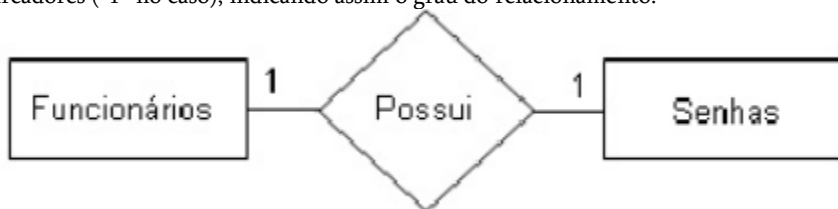
4.4.2.1 Relacionamento Um-para-Um

Esse tipo também é conhecido como **relacionamento 1:1** e define que cada ocorrência de uma entidade somente se associa a uma única ocorrência da outra entidade.



Nem sempre há obrigatoriedade de existir o relacionamento, mas quando houver, ele não pode envolver mais de um par de elementos na associação. Como exemplo pode citar um funcionário que não acessa o sistema, assim não será necessário ter um registro de senha para ele. Para ter um relacionamento **um-para-um**, é preciso que os dois sentidos sejam analisados. Se encontrarmos um único item de uma entidade que se associa a mais de um na outra entidade, não podemos considerar o relacionamento com **um-para-um**.

A figura abaixo mostra o diagrama E-R desse tipo. Note que próximo a cada retângulo que representa as entidades foram adicionados marcadores ("1" no caso), indicando assim o grau do relacionamento.



4.4.2.2 Relacionamento Um-para-Muitos

No segundo tipo de relacionamento (também chamado **relacionamento 1:N**) temos casos em que cada ocorrência de uma entidade **A** pode se relacionar com uma ou mais de uma ocorrência da entidade **B**, porém o contrario não pode acontecer. Temos um bom exemplo do relacionamento entre **Funcionário** e **Filho**.

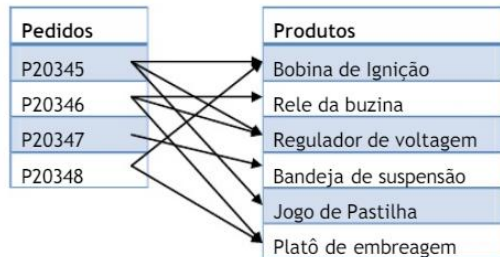


A simbologia utilizada nesse tipo de relacionamento é apresentada a abaixo.

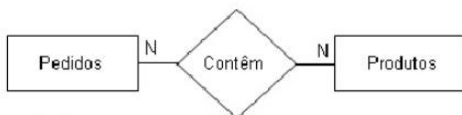


4.4.2.3 Relacionamento Muitos-para-Muitos

Um bom exemplo para entender esse tipo de relacionamento, também denominado **relacionamento N:N**, é o caso de pedidos nos quais constam vários produtos. Um único pedido pode estar associado a vários produtos. Por outro lado, um mesmo produto pode constar em mais de um pedido.





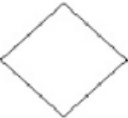




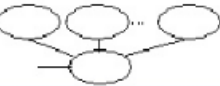




Veja como é representado esse tipo de relacionamento.



4.5 Símbolos de diagramas e convenções de Nomeação

Embora não seja possível assumir como uma regra, é um costume geral adotar para a nomeação das entidades, atributos e relacionamentos nomes escritos no singular com letras maiúsculas.

A tabela abaixo mostra alguns símbolos utilizados em diagramas ER e seus respectivos significados.

Símbolo	Significado
	Entidade
	Entidade fraca
	Relacionamento
	Relacionamento identificado
	Atributo
	Atributo Chave
	Atributo multivalorado
	Atributo composto
	Atributo derivado
	Participação total de E2 em R
	Razão da cardinalidade 1:N para E1:E2 em R
	Restrição estrutural (Min,Max) na participação de E em R