

# **Modelagem e Projeto de Banco de Dados I**

---

*Derivação do Modelo Lógico - Parte I*

# Sumário

---

1 Introdução

2 Conceitos

3 Regras de Derivação

4 Normalização

5 Derivação dos Relacionamentos

6 Derivação das *Generalizações-Especializações*

7 Derivação das *Agregações*

8 Conclusão

# 1. Introdução

---

A obtenção de um modelo lógico a partir de um modelo conceitual pode ser feita pela aplicação de um conjunto de regras bem definidas. Essas regras atuarão, basicamente, em dois grupos distintos de elementos: (i) as estruturas de relacionamento, agregação e especialização de um lado e (ii) as entidades e seus atributos de outro.

Veremos que, através de técnicas bastante objetivas, poderemos, partindo de um modelo conceitual, obter o modelo lógico desejado. As regras a serem apresentadas visarão à geração de um modelo lógico baseado na abordagem relacional.

# 1. Introdução

---

Lembramos que dentro da proposta apresentada em nosso curso, vimos que a geração de um modelo lógico deveria ser antecedida pela obtenção de um modelo conceitual. Esse modelo conceitual deveria ser independente da tecnologia disponível, procurando aproximar-se ao máximo possível do mapeamento fiel do ambiente observado.

A grande vantagem dessa proposta é que, a partir de um modelo conceitual, podemos aplicar regras predefinidas em função da tecnologia a ser empregada e, assim, obter os modelos necessários. Isso fará com que a partir de um mesmo modelo conceitual possamos gerar modelos lógicos para bancos de dados na abordagem REDE, RELACIONAL, OBJETO, etc.

# **1. Introdução**

---

Em síntese, podemos definir que o processo de um modelo lógico, a partir de um modelo conceitual, segue os seguintes passos:

- a) Obter modelo conceitual
- b) Definir tipo de tecnologia (Rede, Relacional, etc)
- c) Aplicar regras de derivação específicas

## **2. Conceitos**

---

Por ser a derivação do modelo lógico fortemente dependente dos conceitos e da tecnologia de implementação a ser utilizada, iremos abordar, de modo bastante objetivo, os principais conceitos envolvidos na abordagem **Relacional**. Contudo, não esgotaremos o assunto, pois seria necessário um trabalho bastante extenso.

O objetivo do nosso curso é a **modelagem de dados baseada no modelo relacional** e não a abordagem relacional por si só.

## **2. Conceitos**

---

O modelo relacional, proposto já na década de 70 por Edgard F. Codd, foi concebido a partir da associação estabelecida entre os conceitos e regras que norteiam a teoria dos conjuntos e as possíveis estruturas de manipulação de dados até então utilizadas.

Codd percebeu, naquele instante, que seria possível aplicar operações conhecidas na matemática, para manuseio de conjuntos, sobre estruturas de dados. Essa abordagem garantiria, a partir daí, um método formal e rigoroso para o tratamento, até então desestruturado, das estruturas de dados.

## **2. Conceitos**

---

A criação, manipulação, combinação, seleção e tantas outras operações realizadas sobre arquivos de dados poderiam passar a ser demonstráveis e comprováveis pelas operações da teoria de conjuntos.

A visão básica estabelecida era de que REGISTROS de um ARQUIVO assemelhavam-se a TUPLAS de uma RELAÇÃO.

Uma RELAÇÃO, por sua vez, assemelhava-se a um ARQUIVO, onde cada uma de suas TUPLAS, contendo elementos de domínios distintos, agregados de forma predefinida, podia ser entendida como um REGISTRO e seus CAMPOS.

## 2. Conceitos

---

Assim, uma Relação R, definida sobre os DOMÍNIOS D1, D2 e D3, seria formada pelas TUPLAS representativas dos DOMÍNIOS D1, D2 e D3 e expressa através de um CONJUNTO:

$R = \{d_{11}, d_{21}, d_{31}; d_{12}, d_{22}, d_{32}; d_{13}, d_{23}, d_{33}; d_{14}, d_{24}, d_{34}; \dots\}$

**segunda tupla**

**(elemento) da relação R**

Por definição, cada elemento da Relação R deve ser único (não repetido na relação) e ter todos os domínios instanciados, não aceitando, p.ex., a ausência de  $d_{21}$  e conseqüentemente a geração de uma ocorrência " $d_{11}, d_{31}$ "

## **2. Conceitos**

---

De acordo com a visão de Codd, o modelo Relacional passou a implementar as estruturas de dados, independentemente de seus tipos, através de uma única visão: TABELAS. O conceito de tabela era a visão tabular ou plana representativa de uma **RELAÇÃO**.

Uma TABELA seguindo a definição de **RELAÇÃO**, seria composta de LINHAS ou TUPLAS, cada uma delas, por sua vez, sendo obtida a partir da instanciamento de domínios predefinidos como formadores da relação. Esses domínios, para o modelo relacional, são tratados como COLUNAS.

## 2. Conceitos

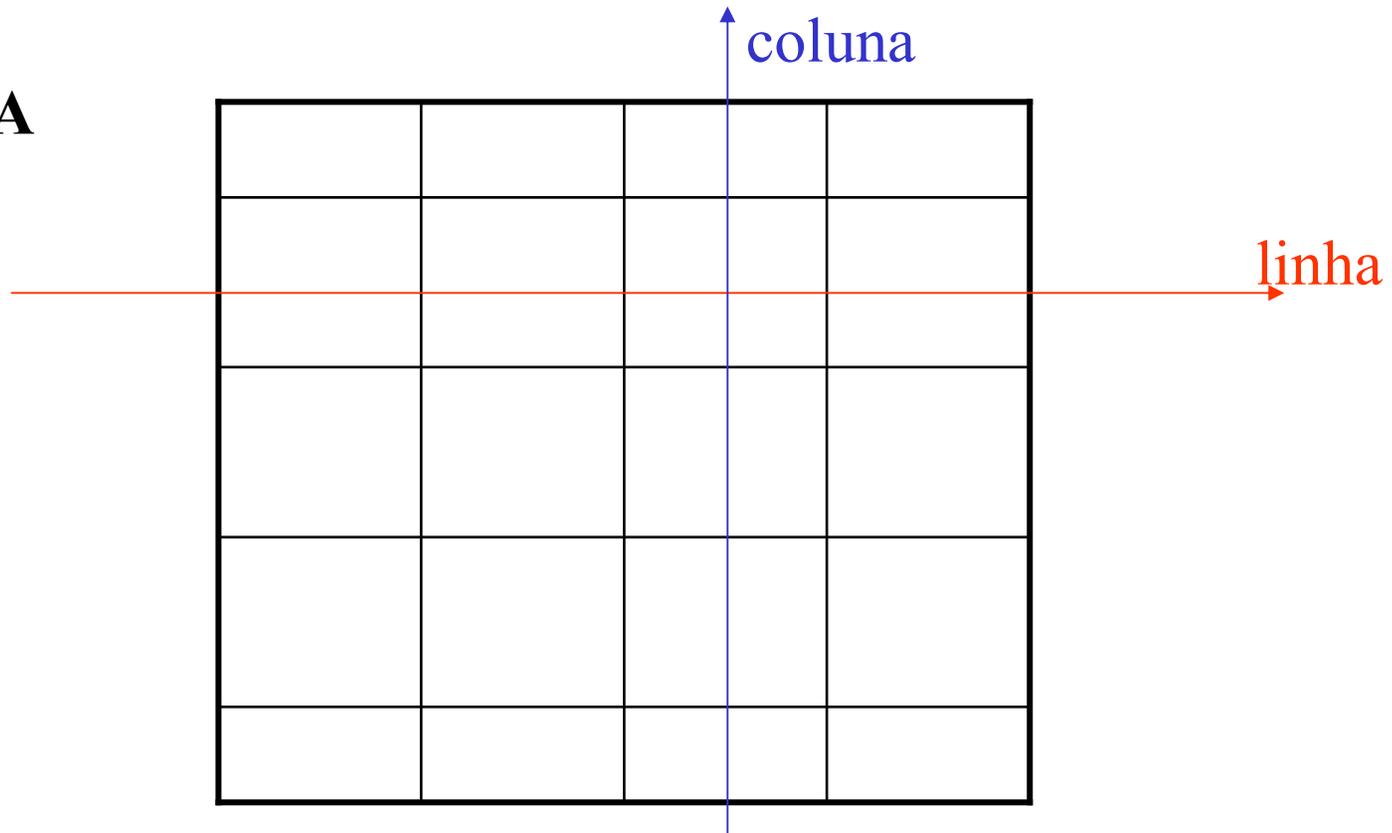
---

<b>DOMÍNIO D1</b>	<b>DOMÍNIO D2</b>	<b>DOMÍNIO D3</b>	<b>DOMÍNIO D4</b>
Nome de Pessoa	Local de Nascimento	Cor de cabelo	Número de CPF
José	Manaus	Ruivo	004
Maria	Salvador	Preto	003
José	Rio	Louro	002

## 2. Conceitos

---

TABELA



No modelo **RELACIONAL**, os dados são representados em forma de **TABELAS** (RELAÇÕES), ou seja, através de **LINHAS** (TUPLAS) e **COLUNAS** (DOMÍNIOS).

## 2. Conceitos

---

### Chave Candidata e Chave Primária (PK - Primary Key)

Entretanto, somente utilizando a estrutura básica de linhas e colunas para construção de tabelas não teremos ainda todos os elementos necessários para a implementação do modelo relacional

Um dos conceitos básicos que definem uma relação é o fato de que “não poderão existir dois elementos iguais dentro de uma relação”.

Mas o que caracteriza um elemento dentro de uma relação? A instanciação de todos os seus domínios!

## 2. Conceitos

---

### Chave Candidata e Chave Primária (PK - Primary Key)

DOMÍNIO D1	DOMÍNIO D2	DOMÍNIO D3	DOMÍNIO D4
Nome de Pessoa	Local de Nascimento	Cor de cabelo	Número de CPF
<b>José</b>	<b>Manaus</b>	<b>Ruivo</b>	<b>004</b>
Maria	Salvador	Preto	003
<b>José</b>	<b>Rio</b>	<b>Louro</b>	<b>002</b>

**José** instanciando 2 vezes o domínio **NOME DE PESSOA**, mas por estar acompanhado por valores diferentes nos demais domínios **NÃO** caracteriza um elemento repetido na relação.

## 2. Conceitos

---

### Chave Candidata e Chave Primária (PK - Primary Key)

Analisar todos os valores de cada um dos domínios pode ser, em alguns casos, bastante complexo e demorado.

Codd estabeleceu uma simplificação que visa a facilitar essa análise. Se, por análise prévia e conhecimento do ambiente que está sendo modelado, formos capazes de identificar domínios nos quais sabemos que a existência de valores instanciados não será repetida, podemos eleger esses domínios, ou colunas, como CANDIDATAS A CHAVE.

A simples análise e comprovação de diferença dessa coluna, comparando seu valor com os demais valores existentes nas demais linhas, fará com que seja possível garantir a não duplicidade de toda uma linha

## 2. Conceitos

---

### Chave Candidata e Chave Primária (PK - Primary Key)

Se analisarmos a coluna Número do CPF veremos que, por definição, não devemos ter duas pessoas com o mesmo CPF. Em outras palavras, se analisarmos nossa tabela e percebermos que duas linhas contêm CPFs diferentes, então obrigatoriamente serão linhas inteiramente diferentes.

Se, contudo, desejarmos admitir que duas pessoas compartilhem o mesmo CPF (marido e mulher, p. ex.), deveremos mudar nossa chave para que, além de CPF, conste também o NOME DA PESSOA como formadora da Chave. Assim, a dupla de colunas deverá ser sempre analisada, e, mesmo com CPFs repetidos, não estaremos caracterizando mais uma duplicidade, desde que os NOMES DE PESSOA sejam diferentes.

## **2. Conceitos**

---

### **Chave Candidata e Chave Primária (PK - Primary Key)**

#### **EXEMPLO**

**TABELA:** Eleitores.

**COLUNAS:** Nome do Eleitor, Data de Cadastramento, Junta Eleitoral, Seção Eleitoral, Número do RG, Número do CPF, Número do Título de Eleitor, Data de Nascimento.

**CHAVES CANDIDATAS:** Número do Título de Eleitor, Número do RG e Número do CPF.

**CHAVE IDENTIFICADORA ou CHAVE PRIMÁRIA (PK – Primary Key):** Número do Título de Eleitor.

Devemos escolher uma das **CHAVES CANDIDATAS** e atribuir a ela a função de **CHAVE IDENTIFICADORA** ou **CHAVE PRIMÁRIA**.

## 2. Conceitos

---

### Chave Estrangeira (FK - Foreign Key)

Como vimos na fase da modelagem conceitual, temos, basicamente, três elementos para representar um ambiente observado: **as entidades, os atributos e os relacionamentos.**

Através de tabelas, portanto, deveremos ser capazes de implementar todos esses elementos. Com os conceitos já vistos até o momento (tabelas, linhas e colunas) poderíamos implementar, facilmente, os dois primeiros elementos: **as entidades e os seus atributos.**

## 2. Conceitos

---

### Chave Estrangeira (FK - Foreign Key)

Para que possamos, entretanto, implementar relacionamentos entre entidades deveremos passar a tratar mais um conceito importante no ambiente relacional: **as chaves estrangeiras**.

#### Características da chave estrangeira:

- é uma chave e, portanto, identifica de modo único uma tupla;
- não está em seu local original (local de origem), mas sim em um local para onde foi migrada (no estrangeiro)
- ela tem um local de origem e lá possui suas características originais (é uma chave primária)

## 2. Conceitos

---

### Chave Estrangeira (FK - Foreign Key)

De forma simples e geral:

→ Uma chave estrangeira nada mais é que uma chave primária de uma tabela que aparece repetida em outra tabela.

Esse é o método para o estabelecimento de relacionamento entre duas tabelas:

→ Basta que se pegue a chave primária de uma das tabelas e a agregue em outra. Essa chave, quando criada em outra tabela, passará a ser uma chave estrangeira.

## 2. Conceitos

---

### Chave Estrangeira (FK - Foreign Key) - Exemplo

Tabela UNIDADE DA FEDERAÇÃO

Chave Primária (PK): SIGLA-UF

Colunas: Nome, Área-m2

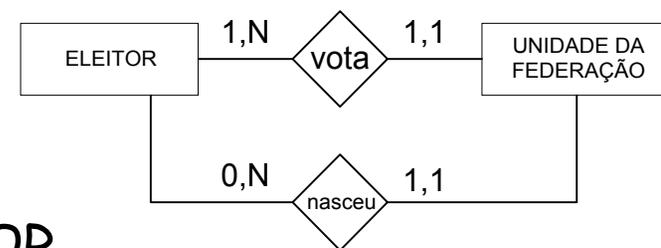
Tabela ELEITOR

Chave Primária (PK): NUM-TÍTULO-ELEITOR

Colunas: CPF, RG, Nome, Estado-Civil

Chave Estrangeira (FK): SIGLA-UF-VOTO

Chave Estrangeira (FK): SIGLA-UF-NASCIMENTO



Pelas chaves estrangeiras podemos relacionar ELEITOR com UNIDADE DA FEDERAÇÃO. Por exemplo, se quisermos identificar o Nome da unidade da federação (Estado ou território) onde alguém vota, poderemos acessar a tabela UNIDADE DA FEDERAÇÃO a partir do conteúdo existente na coluna SIGLA-UF-VOTO

## **2. Conceitos**

---

### **Domínio de Atributos**

O conceito de domínio de um atributo, ou de uma coluna, é fortemente associado ao conceito de domínios de relações.

A idéia principal é que existe um conjunto de valores distintos que podem ser atribuídos a uma coluna.

Entretanto, sabemos que pela diferente natureza de cada uma das colunas poderemos ter diferentes valores sendo válidos às instanciações.

Consideraremos **TRÊS** tipos de domínios:

**2.1 DOMÍNIO DISCRETO**

**2.2 DOMÍNIO CONTÍNUO**

**2.3 DOMÍNIO NULO**

## 2.1 Domínio Discreto

---

### Exemplo

#### Tabela UNIDADE DA FEDERAÇÃO

Chave Primária (PK): SIGLA-UF

Colunas: Nome, Área-m<sup>2</sup>

A coluna SIGLA-UF, que por acaso é a chave primária da tabela, somente pode assumir valores preestabelecidos. Cada unidade da federação quando é criada, recebe um nome, tem sua área delimitada e passa a ser uma SIGLA, ou abreviatura para ser utilizada, por exemplo, no endereçamento postal.

Para a coluna SIGLA-UF, portanto, poderíamos atribuir uma das seguintes siglas: RS, SC, PR, SP, RJ, MG, BA, ... até MA, RO, RR. Isso é o que chamamos de DOMÍNIO DISCRETO: sabemos previamente qual o conjunto permitido de valores, inclusive em suas quantidades.

## 2.2 Domínio Contínuo

---

Exemplo

Tabela UNIDADE DA FEDERAÇÃO

Chave Primária (PK): SIGLA-UF

Colunas: Nome, Área-m<sup>2</sup>

Para a coluna Área-m<sup>2</sup>, qual seria o conjunto permitido de valores? Mesmo se considerarmos a existência de um valor inicial (mínimo) e um valor final (máximo), não podemos, ou não queremos, elencar cada um dos possíveis valores existentes.

Eles caracterizam uma faixa contínua de valores. Isso caracteriza um DOMÍNIO CONTÍNUO.

## 2.3 Domínio Nulo

---

Se estabelecermos uma correlação com a teoria dos conjuntos, um domínio NULO seria representativo de algo como um conjunto vazio.

Assim, alguns atributos, ou colunas, terão seu domínio definido como NULO. O valor aplicável para a instânciação desse atributo é NULO. De forma geral, perceba que não estamos falando em um valor 0 (zero) ou " " (espaço em branco). Um valor NULO não é igual a nada, pois simplesmente não existe.

Consideraremos dois tipos de NULOS:

2.3.1 NULOS INAPLICÁVEIS

2.3.2 NULOS DESCONHECIDOS

## 2.3.1 Nulo Inaplicável

---

São os valores NULOS existentes em uma coluna em função da não aplicabilidade dessa coluna para a linha em questão.

Isso significa que temos uma coluna, representativa de um atributo de um objeto, que em algumas linhas (tuplas) não tem significância. Não pode ser valorado, pois simplesmente não se aplica a esse elemento que está sendo instanciado.

## 2.3.1 Nulo Inaplicável

---

**Exemplo: Ambiente Empresarial**

Tabela: EMPRESA

Colunas: CNPJ, **Inscrição estadual**, Razão Social, Nome fantasia, Endereço, Tipo

Tabela: EMPREGADO

Colunas: RG, Sexo, Data de Nascimento, **Qte Gestações**, Endereço

Alguns domínios são discretos (Tipo, Sexo), outros Contínuos (RG, CNPJ, Data de Nascimento) e outros são **NULOS INAPLICÁVEIS** para algumas instâncias (Inscrição Estadual e Quantidade de Gestações)

## **2.3.1 Nulo Inaplicável**

---

**Exemplo: Ambiente Empresarial**

Tabela: EMPRESA

Colunas: CNPJ, **Inscrição Estadual**, Razão Social, Nome fantasia, Endereço, Tipo

Tabela: EMPREGADO

Colunas: RG, Sexo, Data de Nascimento, **Qte Gestações**, Endereço

**NULOS INAPLICÁVEIS:** Empresas isentas de Inscrição Estadual não terão um valor para a coluna correspondente de Inscrição Estadual; Empregados do sexo masculino não terão um valor para a coluna Qte Gestações.

## 2.3.2 Nulo Desconhecido

---

**Exemplo: Ambiente Empresarial**

Tabela: EMPRESA

Colunas: CNPJ, Inscrição Estadual, Razão Social, Nome fantasia, **Endereço**, Tipo

Tabela: EMPREGADO

Colunas: RG, Sexo, Data de Nascimento, Qte Gestações, **Endereço**

**NULOS DESCONHECIDOS:** Isso ocorre quando um certo valor "ainda" não é conhecido. Significa que a respectiva coluna não teve um valor estabelecido até o momento, mas que ele existe e é aplicável.

## 2.3.2 Nulo Desconhecido

---

**Exemplo: Ambiente Empresarial**

Tabela: EMPRESA

Colunas: CNPJ, Inscrição Estadual, Razão Social, Nome fantasia, **Endereço**, Tipo

Tabela: EMPREGADO

Colunas: RG, Sexo, Data de Nascimento, Qte Gestações, **Endereço**

**NULOS DESCONHECIDOS:** Suponha que um Empregado recém contratado pela Empresa esteja em processo de recrutamento e não saiba informar, no instante de seu cadastramento, o endereço no qual estará habitando. Pode ser atípico, mas é possível.

# 3. Regras de Derivação

---

Considerando a Modelagem RELACIONAL, as regras de derivação do MODELO LÓGICO, a partir do MODELO CONCEITUAL, englobam três atividades distintas:

Normalização das Estruturas de dados

Derivação dos Relacionamentos

\*Derivação de Estruturas Gen-Espec e de Agregação

Sobre essas três atividades, discutiremos nas próximas notas de aula.

\* Optaremos por mostrar esta derivação por último

---

FIM