



Universidade de Brasília

Departamento de Ciência da Computação



Bancos de Dados

CIC0097



Prof. Pedro Garcia Freitas

<https://pedrogarcia.gitlab.io/>

pedro.garcia@unb.br

Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Ciências da Computação



Este conjunto de slides não deve ser utilizado ou republicado sem a expressa permissão do autor.

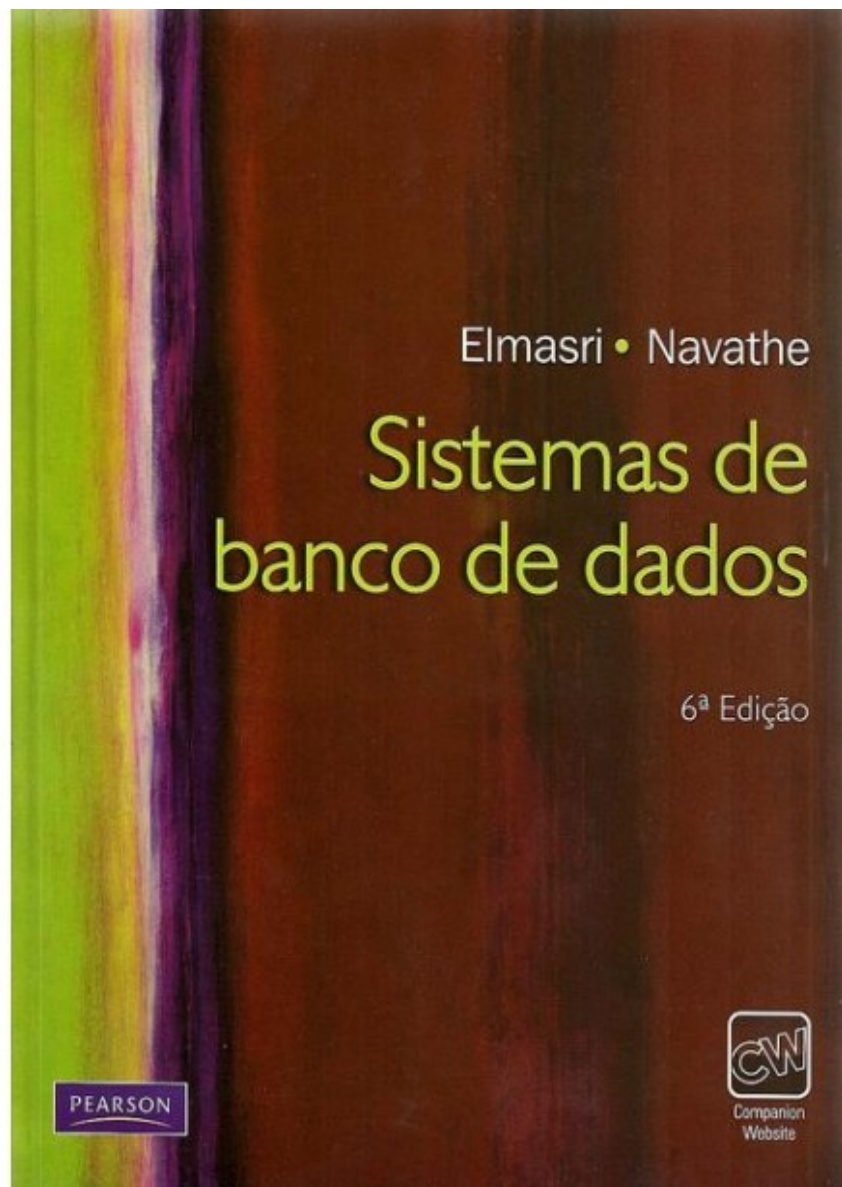
This set of slides should not be used or republished without the author's express permission.



Módulo 17

Projeto de banco de dados: dependências funcionais e normalização – Parte 1 - Diretrizes informais para projeto de banco de dados

**CIC0097/2023.1
T1/T2**



Esta aula se baseia
no Capítulo 15 (SQL
básica) do Elmasri e
Navathe (6^a Edição).



1. Objetivos

Esta aula inicia uma nova fase da nossa disciplina que diz respeito à verificação de qualidade de projetos de banco de dados, começando pelas **diretrizes informais** para que tenhamos a possibilidade de criar um bom projeto de banco de dados.



2. Projeto de Banco de Dados

- Ao desenvolvermos um modelo para um banco de dados, à medida que adquirimos experiência, inevitavelmente estruturamos os atributos das relações de forma coerente e apropriada.
- No entanto, mesmo assim, é essencial comprovar essa adequação e esforçar-se para assegurar a consistência dos dados.
 - I. preservando a informação
 - II. minimizando a redundância



2. Projeto de Banco de Dados

- Os critérios informais para determinar a qualidade de um esquema de relação incluem:
 - I. Garantir que a representação da semântica dos atributos esteja claramente refletida no esquema.
 - II. Reduzir ao mínimo a redundância de informações nas tuplas das relações.

2. Projeto de Banco de Dados

- Os critérios informais para determinar a qualidade de um esquema de relação incluem:

III. Diminuir a frequência de valores NULL nas tuplas.

IV. Eliminar a possibilidade de criação de tuplas indesejadas.

2. Projeto de Banco de Dados

- Assegurar que a semântica dos atributos está claramente representada no esquema.
 - o significado de uma relação está fortemente associado à interpretação dos valores dos atributos em uma tupla;
 - geralmente, quanto melhor o esquema de uma relação mais fácil é atribuir uma semântica a ela;



2. Projeto de Banco de Dados

- Assegurar que a semântica dos atributos está claramente representada no esquema.
 - os esquemas deveriam minimizar a possibilidade de ambiguidades de interpretação;
 - a facilidade em explicar o significado dos atributos em uma relação pode ser tomada como uma medida de qualidade de projeto.

2. Projeto de Banco de Dados

- Vejamos um exemplo de um projeto de banco de dados de baixa qualidade, **e que deve ser evitado.**

Servidor_Depto

<u>cpf</u>	s_nome	s_dt_nasc	endereco	depto_numero	depto_nome	depto_g_ident

Servidor_Projeto

<u>serv_cpf</u>	func_nome	horas	<u>proj_numero</u>	proj_nome	proj_localizacao

2. Projeto de Banco de Dados

- Assegurar que a semântica dos atributos está claramente representada no esquema.

Primeira diretriz: Ao projetar esquemas de relação, garanta que seja simples explicar seu significado. Evite a combinação de atributos de diferentes tipos de entidade e tipos de relacionamento em uma única relação.



2. Projeto de Banco de Dados

- Minimizar a redundância de informação nas tuplas das relações.
- A maneira pela qual os atributos são organizados nas relações:
 - i. tem impactos significativos em relação à utilização de **espaço de armazenamento**.
 - ii. pode resultar em **anomalias de atualização** ao inserir, excluir ou modificar valores.



2. Projeto de Banco de Dados

- Anomalia da **inserção**

Servidor_Depto

<u>cpf</u>	s_nome	s_dt_nasc	endereco	depto_numero	depto_nome	depto_g_ident
1163	Claudia	12/08/1974	SP	5	Engenharia	1163
1164	Jorge	29/01/1986	SP	5	Engenharia	1163
1165	Moacir	08/11/1981	RS	4	Administração	1165

Como inserir um departamento que não possui **servidores**?

2. Projeto de Banco de Dados

- Anomalia da inserção

Servidor_Projeto

<u>serv_cpf</u>	<u>serv_nome</u>	horas	<u>proj_numero</u>	proj_nome	proj_localizacao
1163	Claudia	4	1	ProdutoX	São Paulo
1164	Jorge	6	2	ProdutoY	Sorocaba
1165	Moacir	5	1	ProdutoX	São Paulo
1163	Claudia	6	3	ProdutoZ	São Paulo

A inserção de novos projetos associados a um mesmo funcionário exige a repetição da informação sobre o funcionário.



2. Projeto de Banco de Dados

- Anomalia da inserção

Servidor_Projeto

<u>serv_cpf</u>	<u>serv_nome</u>	horas	<u>proj_numero</u>	proj_nome	proj_localizacao
1163	Claudia	4	1	ProdutoX	São Paulo
1164	Jorge	6	2	ProdutoY	Sorocaba
1165	Moacir	5	1	ProdutoX	São Paulo
1163	Claudia	6	3	ProdutoZ	São Paulo

A associação de mais de um funcionário a um projeto exige a repetição da informação sobre o projeto.



2. Projeto de Banco de Dados

- Anomalia da exclusão

Servidor_Depto

cpf	s_nome	s_dt_nasc	endereco	depto_numero	depto_nome	depto_g_ident
1163	Claudia	12/08/1974	SP	5	Engenharia	1163
1164	Jorge	29/01/1986	SP	5	Engenharia	1163
1165	Moacir	08/11/1981	RS	4	Administração	1165

Se a tupla referente ao funcionário **Moacir** for excluída do sistema, as informações sobre o departamento de **Administração** deixarão de existir no sistema.



2. Projeto de Banco de Dados

- Anomalia da atualização

Servidor_Depto

cpf	s_nome	s_dt_nasc	endereco	depto_numero	depto_nome	depto_g_ident
1163	Claudia	12/08/1974	SP	5	Engenharia	1163
1164	Jorge	29/01/1986	SP	5	Engenharia	1163
1165	Moacir	08/11/1981	RS	4	Administração	1165

Se o nome do departamento 5 mudar, duas tuplas precisarão ser modificadas.



2. Projeto de Banco de Dados

- **Minimizar a redundância de informação nas tuplas das relações.**

Segunda diretriz: ao projetar esquemas de relação, garanta que não seja possível ocorrer nenhuma anomalia durante as operações de inserção, exclusão ou modificação.



2. Projeto de Banco de Dados

- Minimizar a presença de **valores nulos** nas tuplas:
 - Resultam em uma utilização ineficiente do espaço de armazenamento;
 - Provocam dificuldades na compreensão do significado dos atributos;
 - Geram complicações na interpretação do funcionamento de operações de junção e funções agregadas.



2. Projeto de Banco de Dados

- Minimizar a presença de **valores nulos** nas tuplas:

Terceira diretriz: Sempre que possível, evite criar atributos que recebam valores `NULL` frequentemente.



2. Projeto de Banco de Dados

- Eliminar a possibilidade de criação de tuplas espúrias.
 - Tuplas espúrias são tuplas adicionais que são criadas a partir de procedimentos de recuperação de informação mal sucedidos.



2. Projeto de Banco de Dados

- Eliminar a possibilidade de criação de tuplas espúrias.
 - Exemplo: Considerando que o esquema de relação **Servidor_Projeto** é ruim, alguém tentou reprojetar o banco criando duas tabelas para substituir tal relação:
 - **Servidor_Locations**
 - **Servidor_Proj1**



2. Projeto de Banco de Dados

- Eliminar a possibilidade de criação de tuplas espúrias.

Servidor_Projeto

<u>serv_cpf</u>	serv_nome	horas	<u>proj_numero</u>	proj_nome	proj_localizacao
-----------------	-----------	-------	--------------------	-----------	------------------

Servidor_Locations

<u>serv_nome</u>	<u>proj_localizacao</u>
------------------	-------------------------

Servidor_Proj1

<u>serv_cpf</u>	horas	<u>proj_numero</u>	proj_nome	proj_localizacao
-----------------	-------	--------------------	-----------	------------------



2. Projeto de Banco de Dados

Servidor_Locations

<u>serv_nome</u>	<u>proj_localizacao</u>
Claudia	São Paulo
Jorge	Sorocaba
Moacir	São Paulo
Claudia	São Paulo

Servidor_Proj1

<u>serv_cpf</u>	horas	<u>proj_numero</u>	proj_nome	proj_localizacao
1163	4	1	ProdutoX	São Paulo
1164	6	2	ProdutoY	Sorocaba
1165	5	1	ProdutoX	São Paulo
1163	6	3	ProdutoZ	São Paulo



2. Projeto de Banco de Dados

Servidor_Locations

serv

Clau

Jor

Mo

Clau

S

S

- Qual é o resultado da operação

Servidor_Locations |X| **Servidor_PROJ1** ?

- O resultado da execução desse **join** é capaz de recuperar corretamente a informação original?

- É a mesma informação presente na relação **Servidor_PROJ**?

1163

6

3

ProdutoZ

São Paulo



2. Projeto de Banco de Dados

Servidor_Locations |X| Servidor_PROJ1

<u>serv_nome</u>	<u>proj_localizacao</u>	<u>serv_cpf</u>	horas	<u>proj_numero</u>	<u>proj_nome</u>
Claudia	São Paulo	1163	4	1	ProdutoX
Claudia	São Paulo	1165	5	1	ProdutoX
Claudia	São Paulo	1163	6	3	ProdutoZ
Moacir	São Paulo	1163	4	1	ProdutoX
Moacir	São Paulo	1165	5	1	ProdutoX
Moacir	São Paulo	1163	6	3	ProdutoZ
Claudia	São Paulo	1163	4	1	ProdutoX
Claudia	São Paulo	1165	5	1	ProdutoX
Claudia	São Paulo	1163	6	3	ProdutoZ



2. Projeto de Banco de Dados

Servidor_Locations |X| Servidor_PROJ1

serv_nome	proj_localizacao	serv_cpf	horas	proj_numero	proj_nome
Claudia	São Paulo	1163	4	1	ProdutoX
Claudia	São Paulo	1165	5	1	ProdutoX
Claudia	São Paulo	1163	6	3	ProdutoZ
Moacir	São Paulo	1163	4	1	ProdutoX
Moacir	São Paulo	1165	5	1	ProdutoX
Moacir	São Paulo	1163	6	3	ProdutoZ
Claudia	São Paulo	1163	4	1	ProdutoX
Claudia	São Paulo	1165	5	1	ProdutoX
Claudia	São Paulo	1163	6	3	ProdutoZ

Tuplas repetidas



- **Tuplas espúrias** são **registros indesejados** e **incorretos** que podem ser **gerados durante a execução de operações de junção** entre relações de um banco de dados.
- Essas **tuplas adicionais** surgem devido a **problemas na combinação de registros** durante a junção, resultando em resultados **errôneos e inconsistentes**.
- Tuplas espúrias são consideradas indesejadas porque **não representam dados reais e podem causar distorções nas consultas e análises dos dados**.
- É **essencial evitar a ocorrência de tuplas espúrias** para manter a **integridade e a precisão das informações armazenadas no banco de dados**.

2. Projeto de Banco de Dados

- Eliminar a chance de criar tuplas espúrias (falsas ou indesejadas).

Quarta diretriz: Projetar esquemas de relação tal que elas possam ser combinadas em operações de junção com condição de igualdade em atributos que são pares relacionados apropriadamente (chave primária, chave estrangeira) sem que tuplas espúrias sejam geradas.



3. Conclusão

- Resumidamente, temos 4 ***diretrizes informais*** que podem ser usadas como ***medidas para determinar a qualidade*** de projeto do esquema da relação:

1. Garantir que a semântica dos atributos seja clara no esquema.
2. Reduzir a informação redundante nas tuplas.
3. Reduzir os valores **NULL** nas tuplas.
4. Anular a possibilidade de gerar tuplas espúrias.





Dúvidas?

