

# **FACEC**

*F a c u l d a d e*

# Normalização

Professor: Yuri Ferreira

- Revisão aula anterior:
  - SQL Power Architect
    - Como configurar conexão com dataset;
    - Criando modelo Relacional;
    - Persistindo as alterações no BD;
    - Copiando SQL DDL;
  - Exercícios Modelo ER e Mapeamento;

- Conteúdo:
- O que é normalização;
- Normalização: Semântica dos Atributos;
- Redundância de informação;
- Anomalias de atualização;
- Valores nulos nas tuplas;
- Geração de tuplas falsas;
- Dependências Funcionais;

- O que é Normalização?
- A normalização do banco de dados é o processo de **transformações** na **estrutura** de um **BD** que visa eliminar redundâncias e anomalias de inserção, atualização e exclusão (consistência dos dados);
- Aplicando normalização os dados cadastrados no BD ficarão **organizados** e ocuparão **menos espaço físico**;
- Entretanto, **aumenta-se** o número de tabelas e relacionamentos;

- Normalização: Semântica dos Atributos
- A **semântica** de uma relação refere-se ao **significado** da interpretação **dos valores** do atributo em uma tupla (fato ocorrido);
- Em geral quanto mais **fácil for explicar** a semântica da relação, melhor será o projeto do esquema de relação;

- Normalização: Semântica dos Atributos
- Ex: relação **FUNCIONÁRIO**
  - Significado: **cada tupla representa** um funcionário, com valores para o nome, cpf, dt\_nascimento, endereço, chave estrangeira num\_depto que representa relacionamento entre funcionário e departamento;

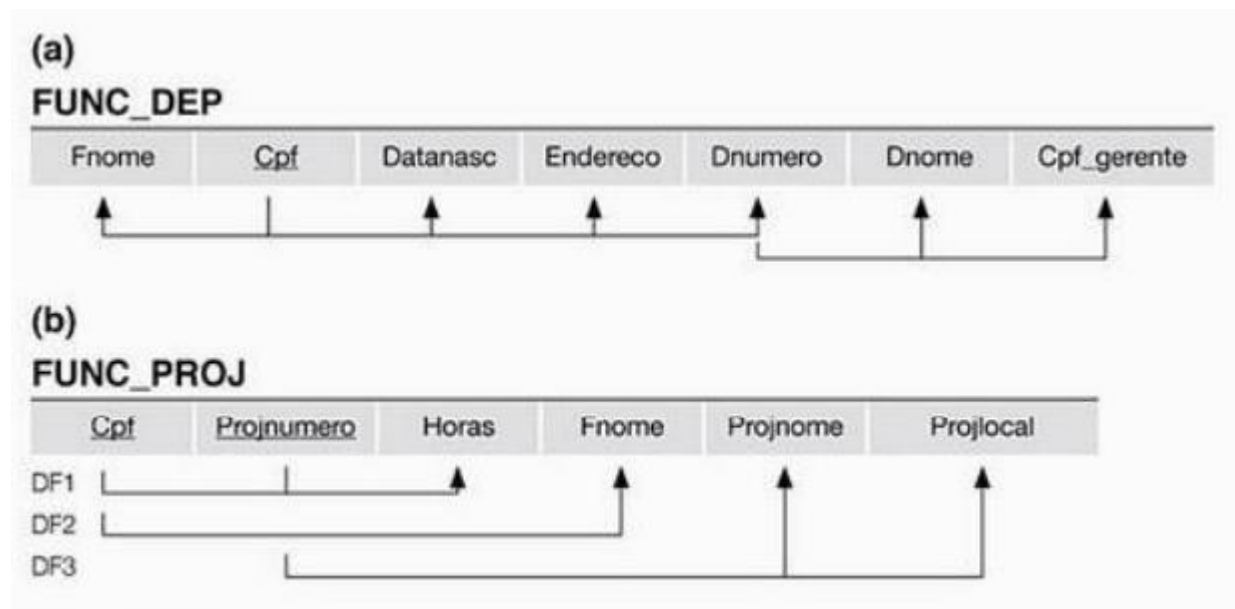
FUNCIONARIO				
PK		FK		
Fnome	<u>Cpf</u>	Datanasc	Endereco	Dnumero

Fonte: Elmasri, 2011.

- Normalização: Semântica dos Atributos
- Em resumo, **Regra 1** - projete um **esquema de relação** de modo que seja fácil de explicar seu **significado**;
- Um esquema de relação que **representa** um tipo de entidade ou relacionamento no **mundo real** tende a ser **claro**;
- Não combine **atributos** de vários tipos de entidade e de **relacionamento** em uma **única relação** (“Tabela”);
  - Para que não haja ambiguidade semântica na relação;



- Normalização: Semântica dos Atributos
- Exemplo violação de qualidade de projeto:



Fonte: Elmasri, 2011.

- Este tipo de relação pode ser utilizado para **visões**, mas não como **relações de base** (tabelas);

- Normalização: Redundância de informação
- Um objetivo do projeto de esquema de um BD é:
  - **minimizar o espaço** de armazenamento usado pelas tabelas;
- Ex: Redundância dos dados do departamento;

FUNC_DEP					Redundância	
Fnome	Cpf	Datanasc	Endereco	Dnumero	Dnome	Cpf_gerente
Silva, João B.	12345678966	09-01-1965	Rua das Flores, 751, São Paulo, SP	5	Pesquisa	33344555587
Wong, Fernando T.	33344555587	08-12-1955	Rua da Lapa, 34, São Paulo, SP	5	Pesquisa	33344555587
Zelaya, Alice J.	99988777767	19-01-1968	Rua Souza Lima, 35, Curitiba, PR	4	Administração	98765432168
Souza, Jennifer S.	98765432168	20-06-1941	Av. Arthur de Lima, 54, Santo André, SP	4	Administração	98765432168
Lima, Ronaldo K.	66688444476	15-09-1962	Rua Rebouças, 65, Piracicaba, SP	5	Pesquisa	33344555587
Leite, Joice A.	45345345376	31-07-1972	Av. Lucas Obes, 74, São Paulo, SP	5	Pesquisa	33344555587
Pereira, André V.	98798798733	29-03-1969	Rua Timbira, 35, São Paulo, SP	4	Administração	98765432168
Brito, Jorge E.	88866555576	10-11-1937	Rua do Horto, 35, São Paulo, SP	1	Matriz	88866555576

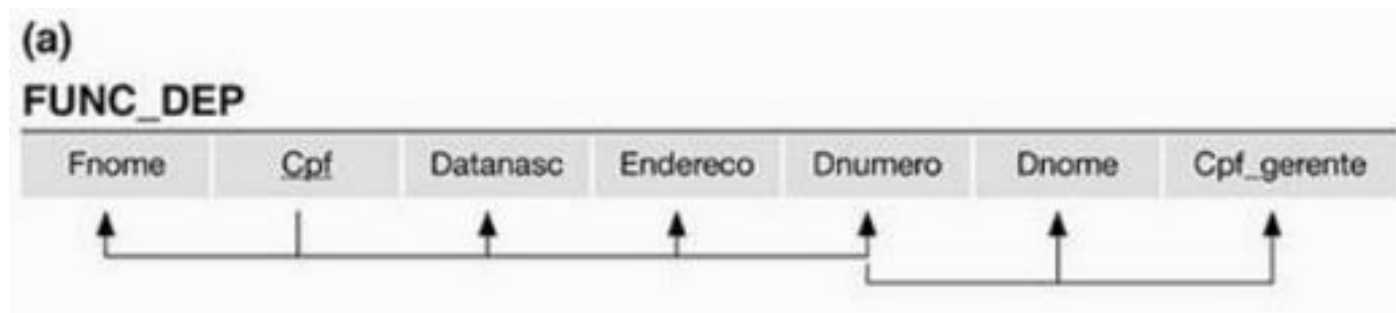
Fonte: Elmasri, 2011.

## ➤ Normalização: Redundância de informação

FUNC_PROJ			Redundância	Redundância	
Cpf	Projnumero	Horas	Fnome	Projnome	Projlocal
12345678966	1	32,5	Silva, João B.	ProdutoX	Santo André
12345678966	2	7,5	Silva, João B.	ProdutoY	Itu
66688444476	3	40,0	Lima, Ronaldo K.	ProdutoZ	São Paulo
45345345376	1	20,0	Leite, Joice A.	ProdutoX	Santo André
45345345376	2	20,0	Leite, Joice A.	ProdutoY	Itu
33344555587	2	10,0	Wong, Fernando T.	ProdutoY	Itu
33344555587	3	10,0	Wong, Fernando T.	ProdutoZ	São Paulo
33344555587	10	10,0	Wong, Fernando T.	Informatização	Mauá
33344555587	20	10,0	Wong, Fernando T.	Reorganização	São Paulo
99988777767	30	30,0	Zelaya, Alice J.	Novosbeneficios	Mauá
99988777767	10	10,0	Zelaya, Alice J.	Informatização	Mauá
98798798733	10	35,0	Pereira, André V.	Informatização	Mauá
98798798733	30	5,0	Pereira, André V.	Novosbeneficios	Mauá
98765432168	30	20,0	Souza, Jennifer S.	Novosbeneficios	Mauá
98765432168	20	15,0	Souza, Jennifer S.	Reorganização	São Paulo
88866555576	20	Null	Brito, Jorge E.	Reorganização	São Paulo

Fonte: Elmasri, 2011.

- Normalização: Anomalias de atualização
- Anomalia de Inserção:
- Inserir informações de forma desnormalizada, implica na manutenção de integridade (coerência) e consistência dos dados;



Fonte: Elmasri, 2011.

- Normalização: Anomalias de atualização
- Anomalia de Inserção:
- Ex: Para **inserir** um **funcionário** inclui-se todos os atributos de departamento OU completa com NULL;



- Os **valores dos atributos** de departamento tem que ser coerentes com os cadastrados na tabela departamento (Difícil de manter);

- Normalização: Anomalias de atualização
- Anomalia de Inserção:
- Ex: Difícil **inserir** um **departamento** que ainda não tenha funcionários;



Fonte: Elmasri, 2011.

- Se funcionário for = NULL, viola restrição de integridade; (chave cpf);

- Normalização: Anomalias de atualização
- Neste projeto, o departamento existe independentemente, se tem funcionários trabalhando para ele;

**FUNCIONARIO**

Fnome	Cpf	Datanasc	Endereco	Dnumero
Silva, Joao B.	12345678966	09-01-1965	Rua das Flores, 751, São Paulo, SP	5
Wong, Fernando T.	33344555587	08-12-1955	Rua da Lapa, 34, São Paulo, SP	5
Zelaya, Alice J.	99988777767	19-01-1968	Rua Souza Lima, 35, Curitiba, PR	4
Souza, Jennifer S.	98765432168	20-06-1941	Av. Arthur de Lima, 54, Santo André, SP	4
Lima, Ronaldo K.	66688444476	15-09-1962	Rua Rebouças, 65, Piracicaba, SP	5
Leite, Joice A.	45345345376	31-07-1972	Av. Lucas Obes, 74, São Paulo, SP	5
Pereira, André V.	98798798733	29-03-1969	Rua Timbira, 35, São Paulo, SP	4
Brito, Jorge E.	88866555576	10-11-1937	Rua do Horto, 35, São Paulo, SP	1

**DEPARTAMENTO**

Dnome	Dnumero	Cpf_gerente
Pesquisa	5	33344555587
Administração	4	98765432168
Matriz	1	88866555576

Fonte: Elmasri, 2011.

- Normalização: Anomalias de atualização
- Anomalia de Exclusão:
- Ex: Se for **excluído** a última tupla de **funcionário** que trabalha para um determ. departamento, as informações do **departamento** serão **perdidas**;
- Esse problema não ocorre quando os dados armazenados são mantidos separadamente (Funcionario, Depart.);



- Normalização: Anomalias de atualização
- Anomalia de Modificação:
- Se for alterado o **gerente** de um **departamento** em uma **tupla**, para manter a integridade, deve se atualizar todas as tuplas de funcionários que tem associação com esse departamento;
- Regra 2: Projete os esquemas de relação tal que nenhuma anomalia de inserção, remoção ou modificação esteja presente nas relações;

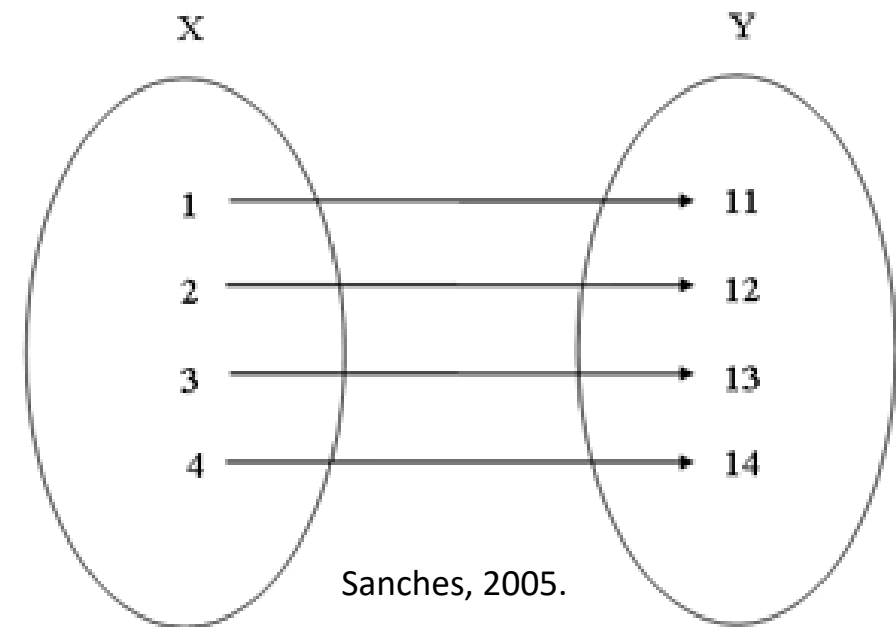
- Normalização: Valores nulos nas tuplas:
- Regra 3: Evite colocar atributos em uma relação cujos valores podem ser **NULL** com **frequência**;
  - Usar o espaço de modo mais eficaz;
  - Evitar junções com valores NULL;
  - são os 2 critérios prioritários que determinam a inclusão das colunas que podem ter NULLs ou que podem ter uma relação separada;

- Normalização: Valores nulos nas tuplas:
- Ex: 15% dos funcionários têm escritórios individuais;
  - Nome\_func|nro\_escritório
- Dessa forma o campo nro\_escritório seria 85% NULL, ocasionando desperdício de espaço;
- Em vez disso uma relação FUNC\_ESCRITORIO pode ser criada para tuplas de funcionários com escritórios individuais;

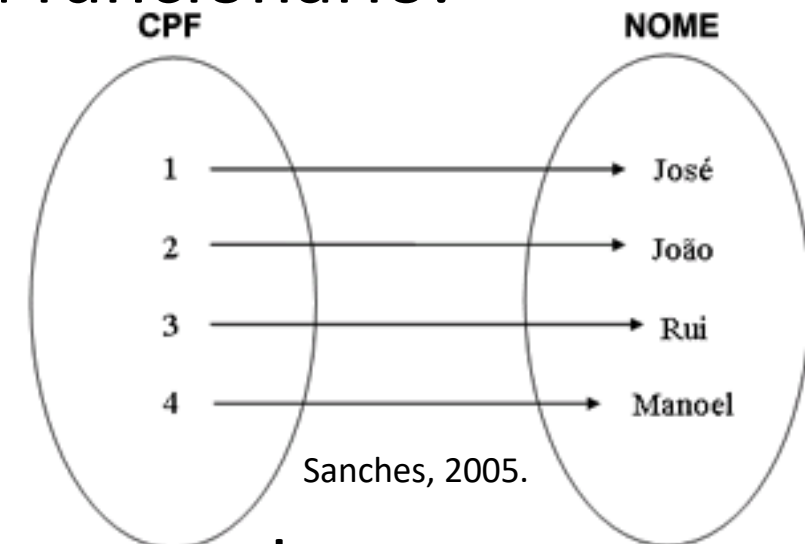
FUNC_ESCRITORIO	
Fcpf	Numero_escritorio

- Normalização: Geração de tuplas falsas:
- Quando fazer um junção entre tabelas se não for utilizado as chaves corretas, pode ser criado linhas falsas que distorcem a informação;
- Regra 4 - Projete o esquema de uma relação tal que **junções** operem em condições de igualdade apenas em **atributos** que são **chaves primárias** ou **chaves estrangeiras**;

- Dependências Funcionais
- Modelo relacional utiliza teoria de funções matemáticas para o conceito de dependência funcional;
- Ex: dependência entre os valores dos conjuntos, pode ser Expressa pela função:  $Y = f(x) = x + 10$



- Dependências Funcionais
- Ex: Dependência entre valores da tabela funcionário:
- Neste caso a função
- **$F(\text{CPF}) = \text{nome}$**
- Ou seja nome é função de CPF;  
isso quer dizer que se tiver o número de CPF encontraremos o nome da pessoa correspondente;
- Notação de dependência: **CPF  $\rightarrow$  NOME**
- Nome depende funcionalmente do CPF;



- Dependências Funcionais
- Regras para encontrar dependências funcionais:
- **1) Separação**
- **$A \rightarrow BC$  então  $A \rightarrow B$  e  $A \rightarrow C$**
- Exemplo:
  - CPF  $\rightarrow$  nome, endereço
  - então CPF  $\rightarrow$  nome e CPF  $\rightarrow$  endereço

- Dependências Funcionais
- Regras para encontrar dependências funcionais:
- **2) Acumulação**
- **A -> B então AC -> B**
- Exemplo:
  - CPF -> endereço então CPF, idade -> endereço
  - Se com o número do CPF encontramos o endereço de uma pessoa, então com este número mais a idade é possível também encontrar o endereço;



- Dependências Funcionais
- Regras para encontrar dependências funcionais:
- **3) Transitividade**
- **$A \rightarrow B$  e  $B \rightarrow C$  então  $A \rightarrow C$**
- Exemplo:
  - CPF  $\rightarrow$  código-cidade e código-cidade  $\rightarrow$  nome-cidade então CPF  $\rightarrow$  nome-cidade;
  - Se com o número do CPF encontramos o código da cidade, e com este código encontramos o nome da cidade, então com o número de CPF é possível encontrar o nome da cidade;

- Dependências Funcionais
- Regras para encontrar dependências funcionais:
- **4) Pseudo-Transitividade**
- **$A \rightarrow B$  e  $BC \rightarrow D$  então  $AC \rightarrow D$**
- Exemplo:
  - CPF  $\rightarrow$  código-funcionario e código-funcionario, mês  $\rightarrow$  salario-funcionario então CPF, mês  $\rightarrow$  salário-funcionario;
  - Se com o número do CPF encontramos o código do funcionario, e com este código mais um certo mês encontramos o salário daquele mês, então com o número do CPF mais um certo mês é possível encontrar o salário que ele recebeu;

## ➤ Formas Normais

- O conceito de normalização foi introduzido por E. F. Codd em 1972;
- Codd criou as três primeiras formas de normalização:
  - primeira forma normal (1NF);
  - segunda forma normal (2NF);
  - terceira forma normal (3NF);

## ➤ Referências:

- SILBERSCHATZ, A.; KORTH, F.; SUDARSHA, S. Database System Concepts. 6. ed. Nova York: MC Graw Hill, 2011.
- ELMASRI, R.; NAVATHE B. Sistemas de banco de dados. 6. Ed. São Paulo, SP: Pearson Addison-Wesley, 2011.
- MORO, M. M. Introdução a Banco de dados. UFMG, 2018.
- Sanches A. R. Fundamentos de armazenamento de manipulação de dados. USP, 2005. Disponível em: <https://www.ime.usp.br/~andrers/aulas/bd2005-1/aula11.html>