

Instituto Federal de Educação, Ciência e Tecnologia da Bahia - IFBA
Departamento de Informática
Integrado / Análise e Desenvolvimento de Sistemas / Licenciatura em Computação

#### Normalização de Relações

André L. R. Madureira <andre.madureira@ifba.edu.br>
Doutorando em Ciência da Computação (UFBA)
Mestre em Ciência da Computação (UFBA)
Engenheiro da Computação (UFBA)

#### Normalização de relações

- Processo de analisar esquemas de relação usando suas dependências funcionais (DFs) e atributos visando:
  - Minimizar redundâncias de dados
  - Minimizar anomalias de atualização
- "Processo de melhoria na qualidade de um banco de dados"
  - Otimização de banco de dados (desempenho e espaço necessário para armazenar os dados)

#### Normalização de relações

- Consiste em testes realizados sobre uma relação para verificar se ela satisfaz uma forma normal
  - Os testes são baseados nas dependências funcionais e atributos de uma relação
- Inicialmente Edgar Codd propôs 3 formas normais em 1972:
  - 1<sup>a</sup> forma normal (1FN)
  - o 2<sup>a</sup> forma normal (2FN)
  - 3<sup>a</sup> forma normal (3FN)

#### Normalização de relações

- Consiste em testes realizados sobre uma relação para verificar se ela satisfaz uma forma normal
  - Os testes são baseados nas dependências funcionais e atributos de uma relação
- Inicialmente Edgar Codd propôs 3 formas normais em 1972:
  - 1<sup>a</sup> forma normal (1FN)
  - 2<sup>a</sup> forma normal (2FN)
  - 3<sup>a</sup> forma normal (3FN)

Relações que não satisfazem à uma forma normal são decompostas em relações menores, que atendam aos testes

#### Forma normal

- Uma relação está em uma forma normal se ela satisfaz TODOS os critérios associados a esta forma normal
  - Uma relação que atende a formas normais de grau maior, também atende às formas normais de grau menor
    - Ex: se uma relação está na 3FN, ela também está na 2FN e na 1FN
  - Grau de normalização: forma normal mais alta da relação

#### 1<sup>a</sup> Forma Normal (1FN)

- Não aceita atributos multivalorados, e nem atributos repetidos
  - Domínio dos atributos inclui apenas:
    - Valores simples, atômicos e indivisíveis
    - Ex: CPF, RG, ano\_nascimento, primeiro\_nome\_mae

Pessoa				
id	<u>cpf</u>	<u>rg</u>	ano_nasc	primeiro_nome_mae
1	415.254.784-00	01354021-67	1992	Carla
2	661.457.745-87	85410234-87	1985	Katia
3	325.471.154-64	74520132-95	1970	Jussara

Pessoa				
id	<u>cpf</u>	ano_nasc	<u>endereco</u>	
1	415.254.784-00	1992	Rua Castelo 74	
2	661.457.745-87	1985	Av ACM 85	
3	325.471.154-64	1970	Rua de Baixo 101	





#### Solução:

Colocar endereço em uma relação à parte, e associar ela a Pessoa

PK	Pessoa	
<u>id</u>	cpf	ano_nasc
1	415.254.784-00	1992
2	661.457.745-87	1985
3	325.471.154-64	1970

Endereço		FK	
rua	numero	id_pessoa	
Rua Castelo	74	1	
Av ACM	85	2	
Rua de Baixo	101	3	

Relação "Pessoa" agora está na 1FN

	Funcionario		
id	cpf	cargo	
1	415.254.784-00	Padeiro	
2	661.457.745-87	Secretaria	
3	325.471.154-64	Padeiro	

	Funcionario		
id	cpf	cargo	
1	415.254.784-00	Padeiro	
2	661.457.745-87	Secretaria	
3	325.471.154-64	Padeiro	

Cargo é um **atributo** repetido

Funcionario		
<u>id</u>	cpf	cargo
1	415.254.784-00	Padeiro
2	661.457.745-87	Secretaria
3	325.471.154-64	Padeiro

Cargo é um **atributo** repetido

Existem pessoas com o mesmo cargo

	Funcionario		
id	cpf	cargo	
1	415.254.784-00	Padeiro	
2	661.457.745-87	Secretaria	
3	325.471.154-64	Padeiro	

Cargo é um **atributo** repetido

Existem pessoas com o mesmo cargo

#### Solução:

Transformar "cargo" em uma relação de "cadastro". Criar uma tabela de junção que associa Funcionario a Cargo

Funcionario		
<u>id</u> cpf		
1	415.254.784-00	
2	661.457.745-87	
3	325.471.154-64	

Cargo		
<u>id</u>	nome	
1	Padeiro	
2	Gerente	
3	Secretaria	

Funcionario		
<u>id</u> cpf		
1	415.254.784-00	
2	661.457.745-87	
3	325.471.154-64	

Cargo		
<u>id</u>	nome	
1	Padeiro	
2	Gerente	
3	Secretaria	

Funcionario_Cargo		
id_func	id_cargo	
1	1	
2	3	
3	1	

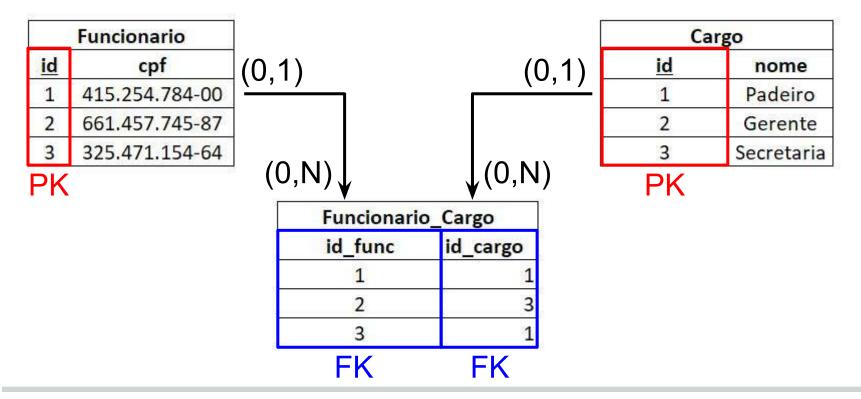
Funcionario		
<u>id</u> cpf		
1	415.254.784-00	
2	661.457.745-87	
3	325.471.154-64	

PK

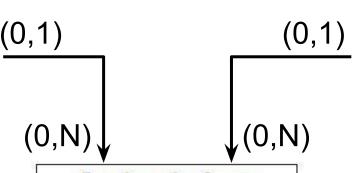
Funcionario_Cargo		
id_func	id_cargo	
1	1	
2	3	
3	1	
FK	FK	

Cargo		
<u>id</u>	nome	
1	Padeiro	
2	Gerente	
3	Secretaria	

PK



Funcionario		
id cpf		
1	415.254.784-00	
2	661.457.745-87	
3	325.471.154-64	



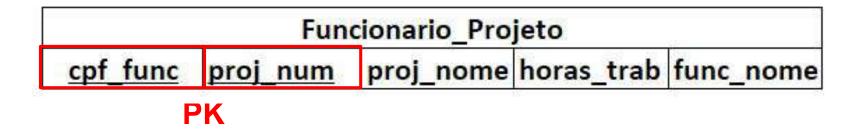
	C	argo
<u>id</u>		nome
	1	Padeiro
2	2	Gerente
	3	Secretaria

Funcionario_Cargo		
id_func	id_cargo	
1	1	
2	3	
3	1	

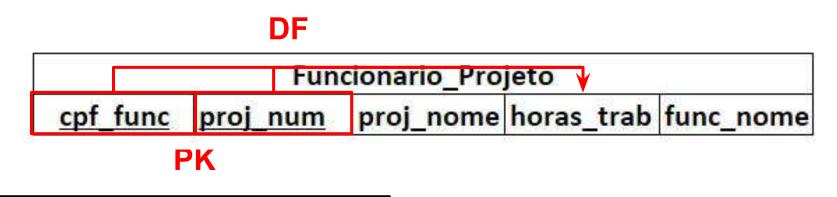
Relação "Funcionario" agora está na 1FN

#### Exercício - Coloque seu projeto lógico na 1FN

- Ajuste o modelo lógico do seu projeto para estar na 1FN
  - Faça os ajustes necessários
  - Considere as boas práticas de bancos de dados
    - Semântica clara (sem ambiguidade)
    - Evite valores NULL
    - Evite tuplas falsas
    - Reduza informações redundantes nas tuplas

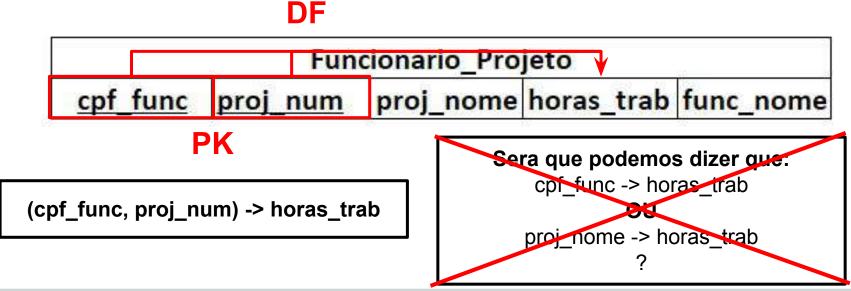


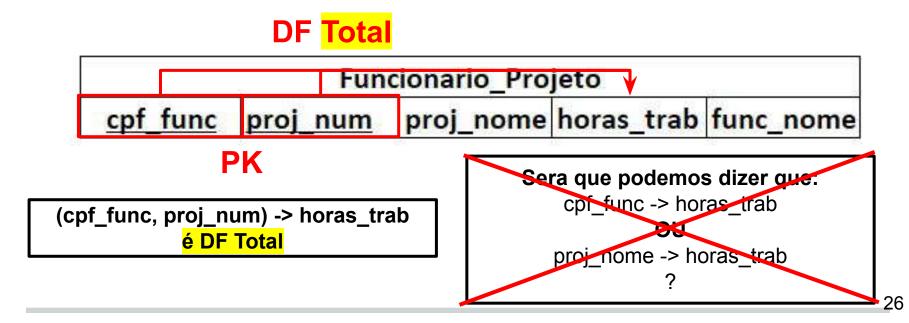
 Uma DF X->Y é uma DF Total se a remoção de qualquer atributo de X significar que a dependência não se mantém mais



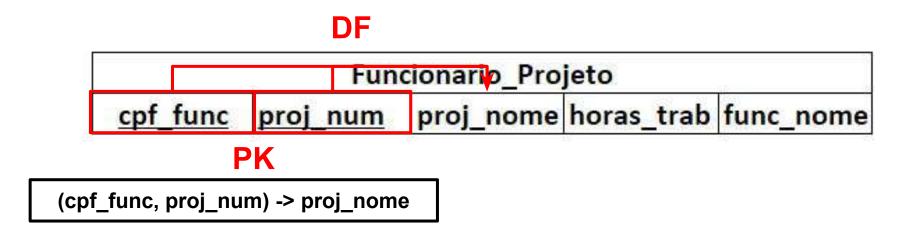
(cpf\_func, proj\_num) -> horas\_trab

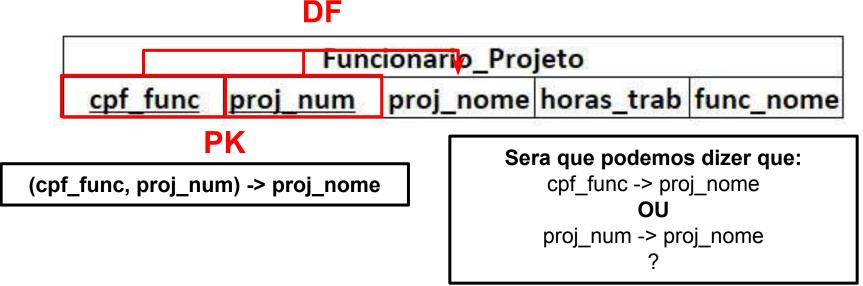


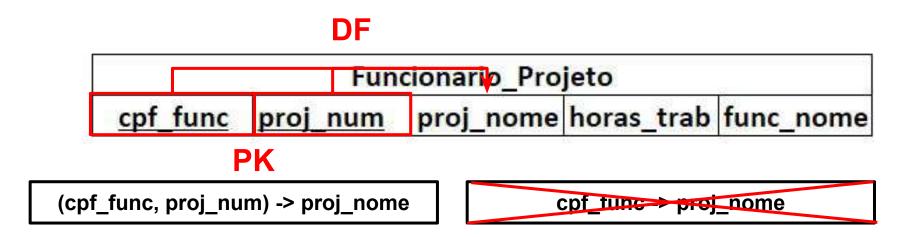


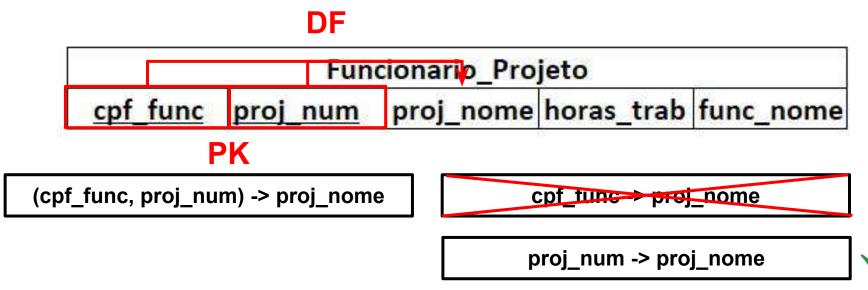


Funcionario_Projeto				
cpf_func	proj_num	proj_nome	horas_trab	func_nome



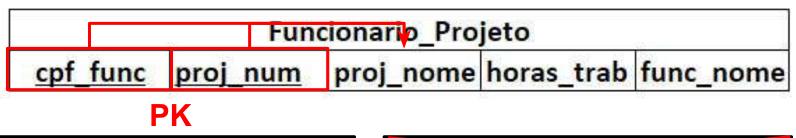






 Uma DF X->Y é uma DF Parcial se algum atributo de X puder ser removido e a DF ainda se mantenha





(cpf\_func, proj\_num) -> proj\_nome é DF Parcial cpf\_tunc > proj\_nome

proj\_num -> proj\_nome

#### 2<sup>a</sup> Forma Normal (2FN)

- Uma relação R está na 2FN, se:
  - Ela também estiver na 1FN
  - Cada atributo n\u00e3o principal A em R for DF total da chave prim\u00e1ria de R

#### 2<sup>a</sup> Forma Normal (2FN)

- Uma relação R está na 2FN, se:
  - Ela também estiver na 1FN

Atributos não pricipais são todos aqueles que não pertencem a chave primária

 Cada atributo n\u00e3o principal A em R for DF total da chave prim\u00e1ria de R

#### 2<sup>a</sup> Forma Normal (2FN)

- Uma relação R está na 2FN, se:
  - Ela também estiver na 1FN

Atributos não pricipais são todos aqueles que não pertencem a chave primária

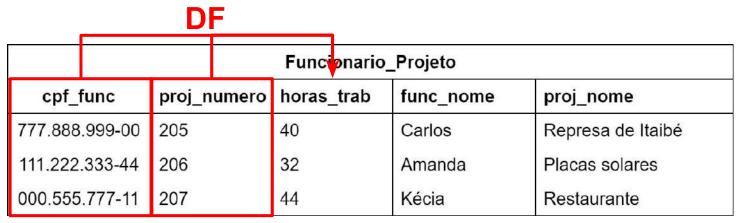
 Cada atributo n\u00e3o principal A em R for DF total da chave prim\u00e1ria de R



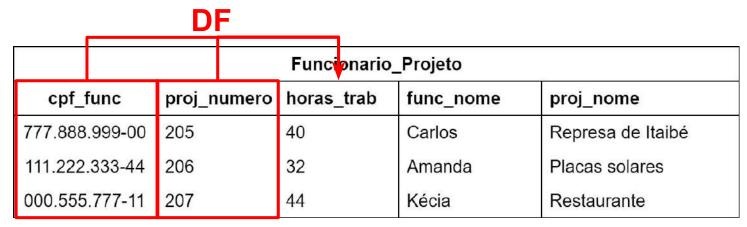
# Exemplo de relação que NÃO está na 2FN

Funcionario_Projeto				
cpf_func	proj_numero	horas_trab	func_nome	proj_nome
777.888.999-00	205	40	Carlos	Represa de Itaibé
111.222.333-44	206	32	Amanda	Placas solares
000.555.777-11	207	44	Kécia	Restaurante





PK



PK

cpf\_func -> horas\_trab

proj\_numero -> horas\_trab

#### **DF Total**

		Funcionario_	_Projeto	
cpf_func	proj_numero	horas_trab	func_nome	proj_nome
777.888.999-00	205	40	Carlos	Represa de Itaibé
111.222.333-44	206	32	Amanda	Placas solares
000.555.777-11	207	44	Kécia	Restaurante

#### PK

cpf func > heras trab



(cpf\_func, proj\_numero) -> horas\_trab
é DF Total

proj numero > horas trab



PK



PK

cpf\_func -> func\_nome

proj\_numero -> func\_nome



PK

cpf\_func -> func\_nome

proj numero > func nome

#### **DF Parcial**

				Funcionario_	_Projeto	1	
cpf_f	unc	proj_n	umero	horas_trab	func_no	ome	proj_nome
777.888	.999-00	205		40	Carlos		Represa de Itaibé
111.222.	333-44	206		32	Amanda	Ĺ	Placas solares
000.555	.777-11	207		44	Kécia		Restaurante

PK

cpf\_func -> func\_nome



proj numero > func nome

#### **DF Parcial** Funcionario\_Projeto cpf func proj numero horas trab func nome proj nome 777.888.999-00 205 40 Carlos Represa de Itaibé 111.222.333-44 32 206 Amanda Placas solares 000.555.777-11 207 44 Kécia Restaurante PK

Como existem atributos que **dependem parcialmente** da chave primária (*cpf\_func*, *proj\_numero*), a relação NÃO ESTÁ na 2FN

#### **DF Total**

Funcionario Projeto						
cpf_func						
777.888.999-00	205	40				
111.222.333-44	206	32				
000.555.777-11	207	44				

PK

cpf func > heras trab

proj num<del>ero ≻ horas</del> trab



(cpf\_func, proj\_numero) -> horas\_trab é **DF Total** 

Funcionario_Projeto						
cpf_func   proj_numero   horas_trak						
777.888.999-00	205	40				
111.222.333-44	206	32				
000.555.777-11	207	44				

#### **DF Total**

uncionario					
<u>cpf</u>	nome				
777.888.999-00	Carlos				
111.222.333-44	Amanda				
000.555.777-11	Kécia				

PK

Não é preciso testar se **cpf** -> **nome** é DF Total. Isto se deve ao fato de termos apenas um atributo principal (*cpf*).

Funcionario_Projeto						
cpf_func	proj_numero	horas_trab				
777.888.999-00	205	40				
111.222.333-44	206	32				
000.555.777-11	207	44				

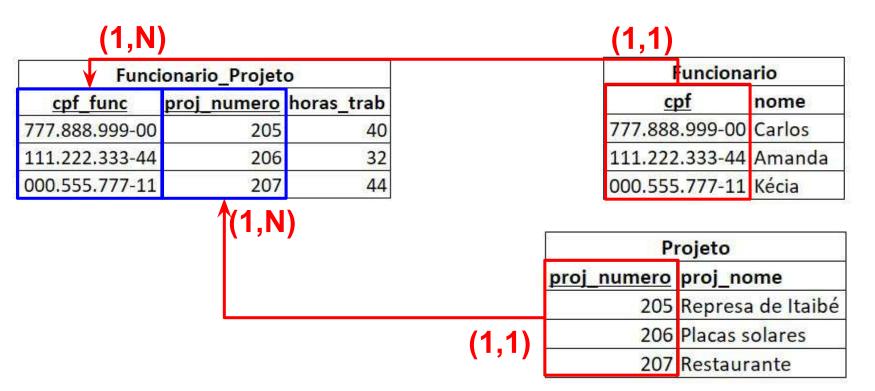
Funcionario					
cpf nome					
777.888.999-00	Carlos				
111.222.333-44	Amanda				
000.555.777-11	Kécia				

#### **DF Total**

Projeto					
<u>proj_numero</u> proj_nome					
205	Represa de Itaibé				
206	Placas solares				
207	Restaurante				



FK Funci	ionario_Projet	o			unciona	rio
cpf_func	proj_numero	horas_trab		<u>c</u>	pf	nome
777.888.999-00	205	40		777.888	3.999-00	Carlos
111.222.333-44	206	32		111.222	2.333-44	Amanda
000.555.777-11	207	44		000.555	5.777-11	Kécia
	FK↑			P	K	
				P	rojeto	
			proj	numero	proj_no	me
				205	Represa	a de Itaibé
				206	Placas s	olares
				207	Restaur	ante
			28	PK	(/ <u></u>	



### Exercício - Coloque seu projeto lógico na 2FN

- Ajuste o modelo lógico do seu projeto para estar na 2FN
  - Faça os ajustes necessários
  - Considere as boas práticas de bancos de dados
    - Semântica clara (sem ambiguidade)
    - Evite valores NULL
    - Evite tuplas falsas
    - Reduza informações redundantes nas tuplas

- Uma DF X -> Y é uma DF transitiva se existe:
  - X -> W
  - $\circ$  W -> Y
- Onde W é um atributo não principal ("W não é chave primária")

Funcionario_Departamento							
cpf_func nome_func data_nasc num_depto cpf_gerente							
777.888.999-00	Carlos	1985-05-02	100	504.231.854-78			
111.222.333-44	Jussara	1995-07-05	101	854.632.140-98			
000.555.777-11	Maria	1970-02-20	100	504.231.854-78			

- Uma DF X -> Y é uma DF transitiva se existe:
  - X -> W
  - $\circ$  W -> Y
- Onde W é um atributo não principal ("W não é chave primária")

#### DF

Funcionario_Departamen o						
cpf_func	nome_func	data_nasc	num_depto	cpf_gerente		
777.888.999-00	Carlos	1985-05-02	100	504.231.854-78		
111.222.333-44	Jussara	1995-07-05	101	854.632.140-98		
000.555.777-11	Maria	1970-02-20	100	504.231.854-78		



- Uma DF X -> Y é uma DF transitiva se existe:
  - X -> W
  - $\circ$  W -> Y
- Onde W é um atributo não principal ("W não é chave primária")

#### DF

Funcionario_Departamen o							
cpf_func nome_func data_nasc num_depto cpf_gerente							
777.888.999-00	Carlos	1985-05-02	100	504.231.854-78			
111.222.333-44	Jussara	1995-07-05	101	854.632.140-98			
000.555.777-11	Maria	1970-02-20	100	504.231.854-78			

PK



#### DF

Funcionario_Departamen o					
cpf_func	nome_func	data_nasc	num_depto	cpf_gerente	
777.888.999-00	Carlos	1985-05-02	100	504.231.854-78	
111.222.333-44	Jussara	1995-07-05	101	854.632.140-98	
000.555.777-11	Maria	1970-02-20	100	504.231.854-78	

PK

DF

#### Temos:

cpf\_func -> num\_depto
num\_depto -> cpf\_gerente

#### DF

Funcionario_Departamento					
cpf_func	nome_func	data_nasc	num_depto	cpf_gerente	
777.888.999-00	Carlos	1985-05-02	100	504.231.854-78	
111.222.333-44	Jussara	1995-07-05	101	854.632.140-98	
000.555.777-11	Maria	1970-02-20	100	504.231.854-78	

PK

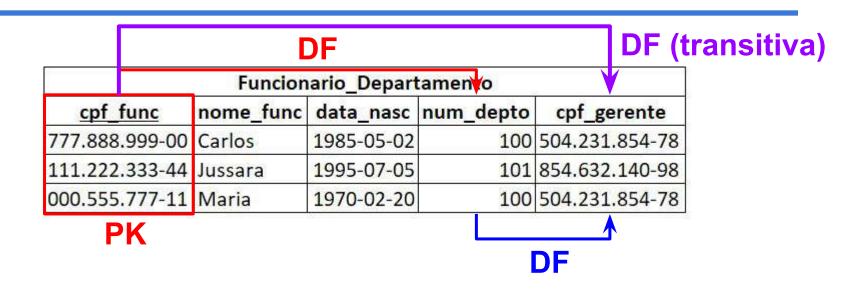
DF

#### Temos:

cpf\_func -> num\_depto
num\_depto -> cpf\_gerente



cpf\_func -> cpf\_gerente
 existe e é DF transitiva



#### Temos:

cpf\_func -> num\_depto
num\_depto -> cpf\_gerente



cpf\_func -> cpf\_gerente
 existe e é DF transitiva

#### 3<sup>a</sup> Forma Normal (3FN)

- Uma relação R está na 3FN, se:
  - Ela também estiver na 2FN
  - Nenhum atributo não principal (que não pertence a chave primária) de R for DF transitiva da chave primária
    - Isto é, atributos não devem depender de outros atributos não principais

		Funcion	ario_Depart	tament	0		
cpf	func	nome_func	data_nasc	num	lepto	cpf_ge	erente
777.888	3.999-00	Carlos	198-25-12		100	504.231	.854-78
111.222	2.333-44	Jussara	1995-07-05		101	854.632	.140-98
000,555	777-11	Maria	1970-02-20		100	504.231	.054-78
PK			2	i			
		DF (tra	nsitiva)			DF	

Temos as seguintes DFs:

cpf\_func -> num\_depto -> cpf\_gerente

Funcionario_Departamento					
cpf_func	nome_func	data_nasc	num_depto	cpf_gerente	
777.888.999-00	Carlos	1985-05-02	100	504.231.854-78	
111.222.333-44	Jussara	1995-07-05	101	854.632.140-98	
000.555.777-11	Maria	1970-02-20	100	504.231.854-78	

	Funcion	ario_Depart	amento	
cpf_func	nome_func	data_nasc	num_depto	cpf_gerente
777.888.999-00	Carlos	1985-05-02	100	504.231.854-78
111.222.333-44	Jussara	1995-07-05	101	854.632.140-98
000.555.777-11	Maria	1970-02-20	100	504.231.854-78

Pertence a tabela Funcionário

Pertence a tabela Departamento

Funcionario			
<u>cpf</u>	nome	data_nasc	
777.888.999-00	Carlos	1985-05-02	
111.222.333-44	Jussara	1995-07-05	
000.555.777-11	Maria	1970-02-20	

Departamento		
numero	cpf_gerente	
100	504.231.854-78	
101	854.632.140-98	

Funcionario			
<u>cpf</u>	nome	data_nasc	
777.888.999-00	Carlos	1985-05-02	
111.222.333-44	Jussara	1995-07-05	
000.555.777-11	Maria	1970-02-20	

Departamento		
numero	cpf_gerente	
100	504.231.854-78	
101	854.632.140-98	

Funcionario_Departamento		
cpf_func	num_depto	
777.888.999-00	100	
111.222.333-44	101	
000.555.777-11	100	

Fun	cionario	0	De	epartamento
<u>cpf</u>	nome	data_nasc	numer	o cpf_gerente
777.888.999-00	Carlos	1985-05-02	100	504.231.854-78
111.222.333-44	Jussara	a 1995-07-05	101	854.632.140-98
000.555.777-11	Maria	1970-02-20	13	(1,1)
(1,1) (1,N)		Funcionario_De	partamento	
		cpf_func	num_depto	
		777.888.999-00	100	<del></del>
		111.222.333-44	101	(1,N)
	9	000.555.777-11	100	

#### Resumo das Formas Normais

- Primeira Forma Normal (1FN)
  - Não há atributos multivalorados, repetidos, nem compostos
- Segunda Forma Normal (2FN)
  - Não há **DF parcial entre** atributos chaves e não chaves
- Terceira Forma Normal (3FN)
  - Não há **DF transitiva** entre atributos não chaves

### Exercício - Coloque seu projeto lógico na 3FN

- Ajuste o modelo lógico do seu projeto para estar na 3FN
  - Faça os ajustes necessários
  - Considere as boas práticas de bancos de dados
    - Semântica clara (sem ambiguidade)
    - Evite valores NULL
    - Evite tuplas falsas
    - Reduza informações redundantes nas tuplas

### Referencial Bibliográfico

 KORTH, H.; SILBERSCHATZ, A.; SUDARSHAN, S.
 Sistemas de bancos de dados. 5. ed. Rio de Janeiro: Ed. Campus, 2006.

 DATE, C. J. Introdução a sistemas de bancos de dados. Rio de Janeiro: Ed. Campus, 2004. Tradução da 8ª edição americana.