

Teoria da Normalização Dependências Funcionais

Prof. Antonio Guardado

Normalização de Dados

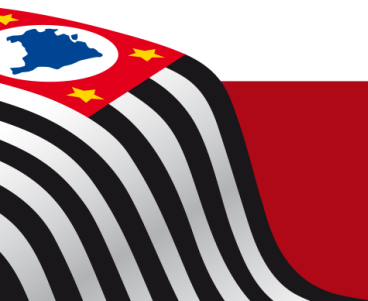
É um **processo formal**, passo a passo, de **análise dos atributos** de uma relação

1. Objetivos

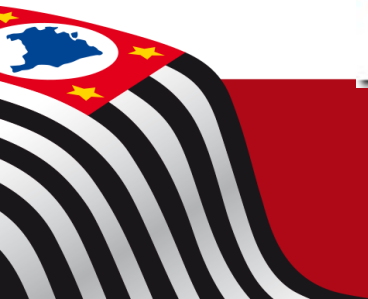
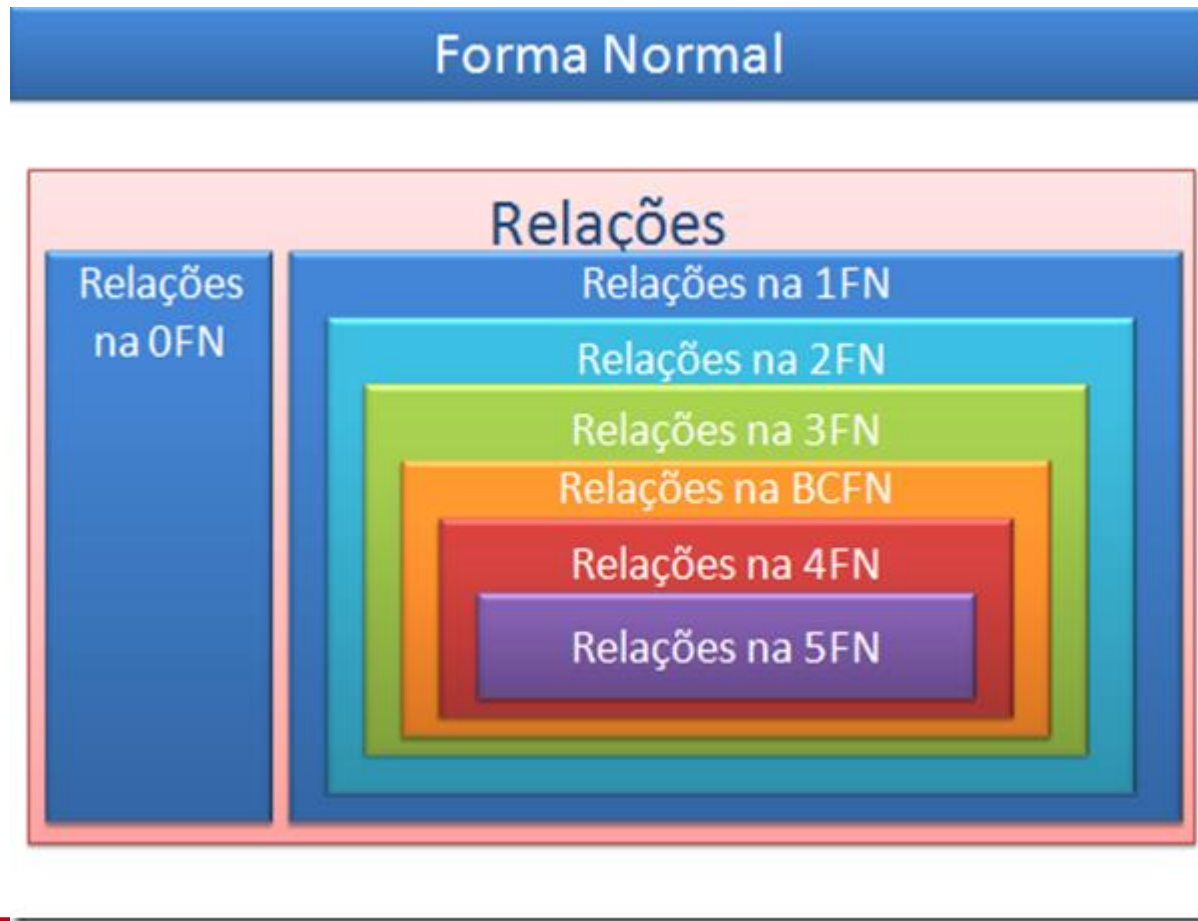
- eliminar redundâncias
- evitar anomalias de atualização, exclusão e atualização

2. Baseado no conceito de formas normais

- processo sistemático de geração de tabelas
- estar em uma forma normal significa atender uma regra
- 5 formas normais: 1, 2 e 3 mais aplicadas
- Estar na forma normal n significa estar na forma normal $(n-1)$
- Utiliza o conceito de dependência funcional



Formas Normais



Anomalias de Atualização

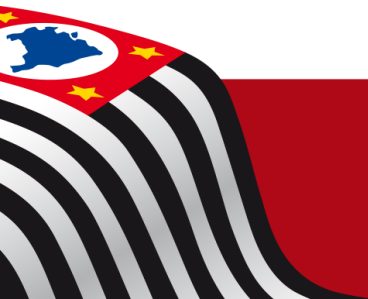
Inclusão

Exclusão

Atualização

Ex: Quais problemas são decorrentes da relação Vendas?

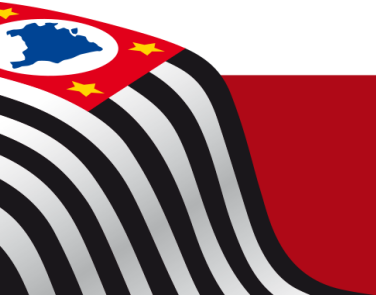
nomeCli	<u>CPF</u>	endereco	fone	<u>codProd</u>	nomeProd	Vunit	qtde	total
Zé	111	ABC	9923123	A	Lápis	0,50	2	1,00
Ana	222	XYZ	9111456	B	Caneta	1,00	3	3,00
Zé	111	ABC	9912123	C	Régua	1,00	2	2,00
Pedro	444	KZZ	Null	A	Lápis	0,50	20	10,00



Anomalias de Atualização

- Modificação/Atualização:
 - uma mudança na descrição do produto A requer várias mudanças
- inconsistência:
 - não há nada no projeto impedindo que o produto A tenha duas ou mais descrições diferentes no BD

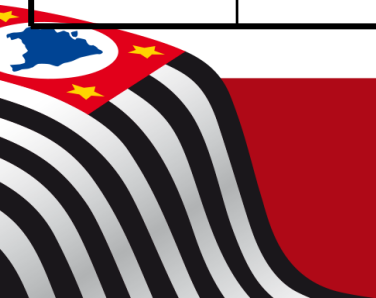
nomeCli	<u>CPF</u>	endereco	fone	<u>codProd</u>	nomeProd	Vunit	qtde	total
Zé	111	ABC	9923123	A	Lápis	0,50	2	1,00
Ana	222	XYZ	9111456	B	Caneta	1,00	3	3,00
Zé	111	ABC	9912123	C	Régua	1,00	2	2,00
Pedro	444	KZZ	Null	A	Lápis	0,50	20	10,00



Anomalias de Inclusão e Exclusão

- Inclusão:
 - redundância do cliente e do produto
- Exclusão:
 - se o cliente ANA fosse eliminado seria perdida a informação de que o produto B é chamado *caneta e custa R\$ 1,00*

nomeCli	<u>CPF</u>	endereco	fone	<u>codProd</u>	nomeProd	Vunit	qtde	total
Zé	111	ABC	9923123	A	Lápis	0,50	2	1,00
Ana	222	XYZ	9111456	B	Caneta	1,00	3	3,00
Zé	111	ABC	9912123	C	Régua	1,00	2	2,00
Pedro	444	KZZ	Null	A	Lápis	0,50	20	10,00

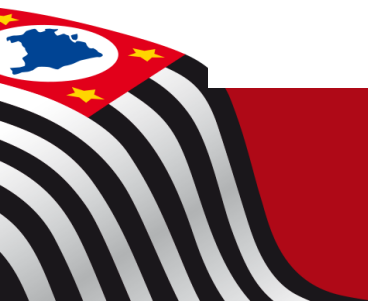
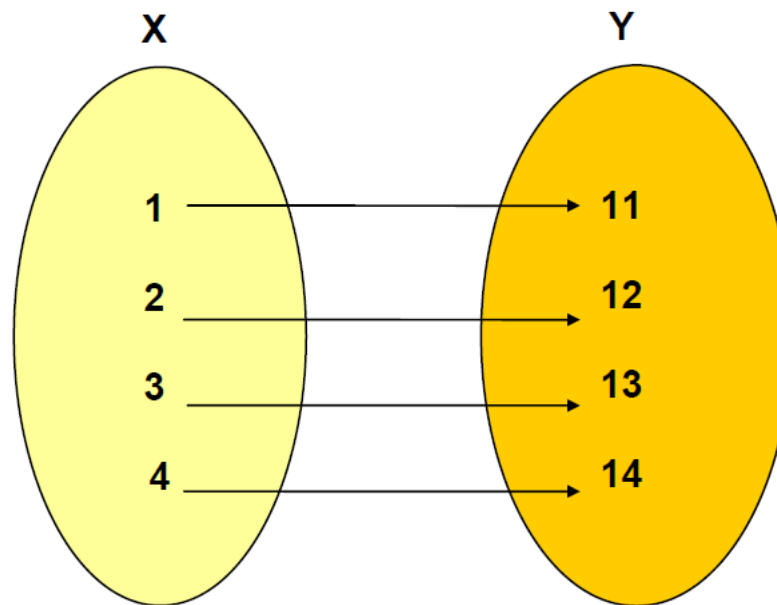


Dependências Funcionais

O Modelo Relacional tomou emprestado da teoria de funções da matemática o conceito de dependência funcional.

Iremos utilizar então a teoria de funções para explicar a dependência funcional do Modelo Relacional.

Considerando os seguintes conjuntos :

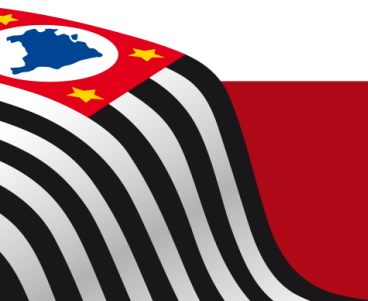
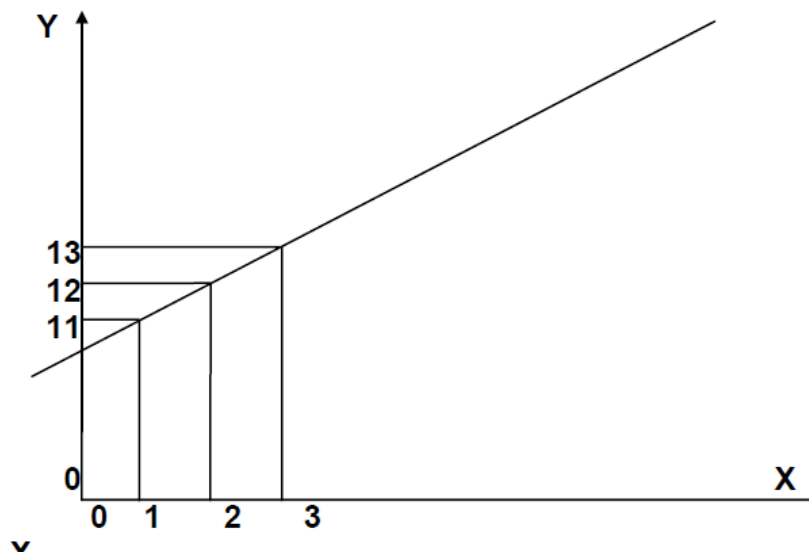


Dependências Funcionais

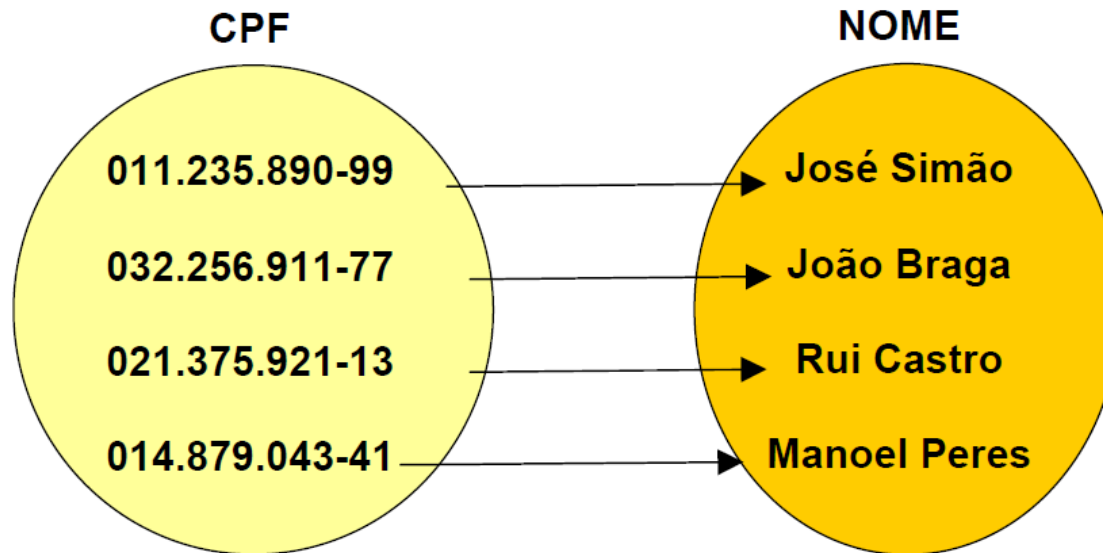
Observe que existe uma dependência entre os valores dos conjuntos, que pode ser expressa pela função

$$\begin{aligned} f(x) &= x + 10, \\ \text{ou seja,} \\ y &\text{ é função de } x, \\ \text{ou seja,} \\ y &= f(x) = x + 10. \end{aligned}$$

Esta função pode também ser expressa através do gráfico ao lado.



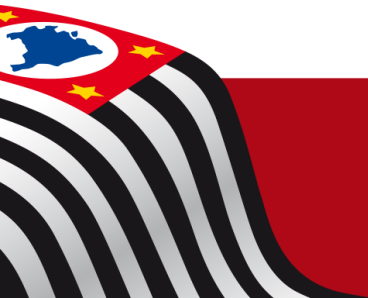
Dependências Funcionais



$f(\text{CPF}) = \text{nome}$

nome é função do CPF

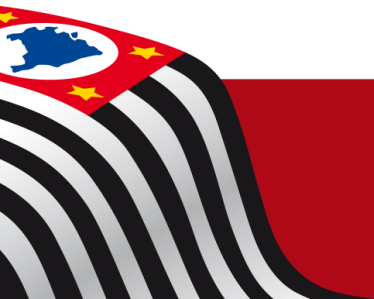
$\text{CPF} \rightarrow \text{NOME}$



Definição DF em Banco de Dados

Em Banco de dados define-se dependência funcional da seguinte forma (simplificada) :

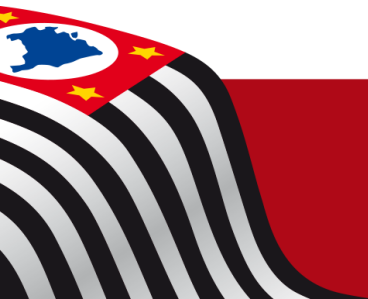
Dada uma relação r , o atributo Y de r é funcionalmente dependente do atributo X de r **se somente se** sempre que duas tuplas de r combinarem em seus valores de X elas também combinarem no valor de Y .



Definição Formal em Banco de Dados

Formalmente :

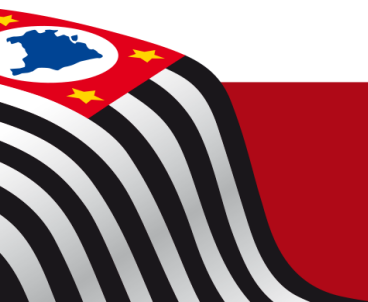
Seja r uma relação e X e Y subconjuntos arbitrários do conjunto de atributos de r : $X, Y \subseteq R$, $R = \{ A_1, \dots, A_n \}$ onde R é a relação esquema para a qual $r = r(R)$. Então, dizemos que Y é funcionalmente dependente de X , expresso como $X \rightarrow Y$, ou ainda $A_1 \dots A_j \rightarrow A_k \dots A_n$ (A_j e A_k não necessariamente distintos), se e somente se cada valor de X em r tiver associado a ele precisamente um valor Y em r . Em outras palavras, sempre que duas tuplas de r concordarem sobre seu valor X , elas também concordarão sobre seu valor Y .



Exemplo

Exemplo : Código → Salário

	Código	Salário
	E1		10
	E3		10
	E1		10
	E2		5
	E3		10
	E2		5
	E1		10



Regras para encontrar Dependências Funcionais

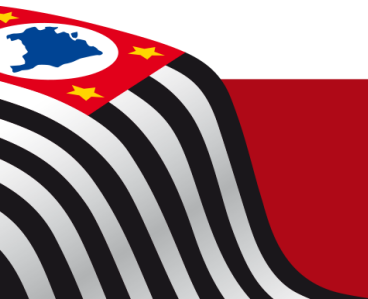
1. Separação

$A \rightarrow BC$ então $A \rightarrow B$
e $A \rightarrow C$

Exemplo :

$CPF \rightarrow \text{nome, endereço}$ então $CPF \rightarrow \text{nome}$
e $CPF \rightarrow \text{endereço}$

Se com um número de CPF eu encontro o nome e o endereço de uma pessoa, então com este mesmo número eu posso encontrar apenas o nome, e com este mesmo número eu posso encontrar apenas o endereço.



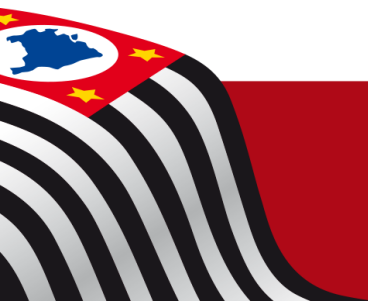
2. Acumulação

$A \rightarrow B$ então $AC \rightarrow B$

Exemplo :

$\text{CPF} \rightarrow \text{endereço}$ então $\text{CPF, sexo} \rightarrow \text{endereço}$

Se com um número de CPF eu encontro o endereço de uma pessoa, então com este mesmo número mais o sexo da pessoa eu posso encontrar o endereço também.



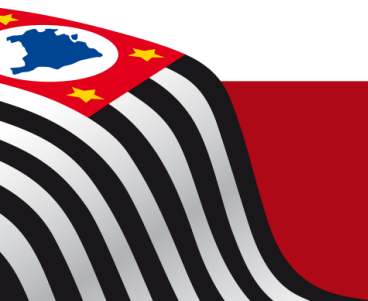
3. Transitividade

$A \rightarrow B$ 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 um número de CPF eu encontro o código da cidade de uma pessoa, e com o código da cidade eu encontro o nome da cidade, então com o número do CPF eu posso encontrar o nome da cidade.



4. Pseudo-Transitividade

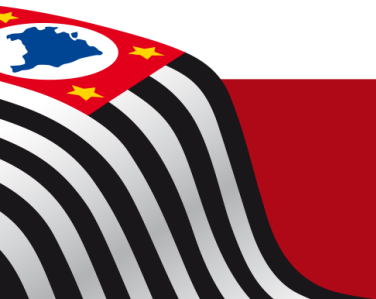
$A \rightarrow B$ e $BC \rightarrow D$ então $AC \rightarrow D$

Exemplo :

CPF \rightarrow código-funcionário e
código-funcionário, mês \rightarrow salário-funcionário
então

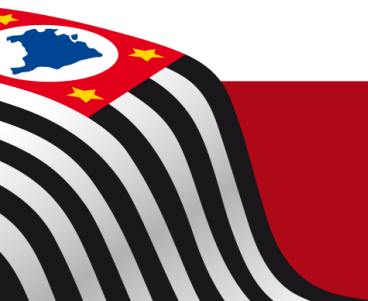
CPF, mês \rightarrow salário-funcionário

Se com um número de CPF eu encontro o código do funcionário, e com o código do funcionário mais um certo mês eu encontro o salário que ele recebeu naquele mês, então com o número do CPF mais um certo mês eu posso encontrar o salário que ele recebeu naquele mês.



Processo de Normalização

- inicia com uma relação ou coleção de relações
- produz uma nova coleção de relações:
 - equivalente a coleção original (representa a mesma informação)
 - livre de problemas
- Propriedades a serem alcançadas:
 - Minimização da redundância
 - Minimização de anomalias de inserção, exclusão e atualização de dados
 - Ex.: *Florianópolis, Florianópolis, Floripa, Flópolis*
- **Significado: as novas relações estarão, pelo menos na 3FN**



Processo de Normalização

