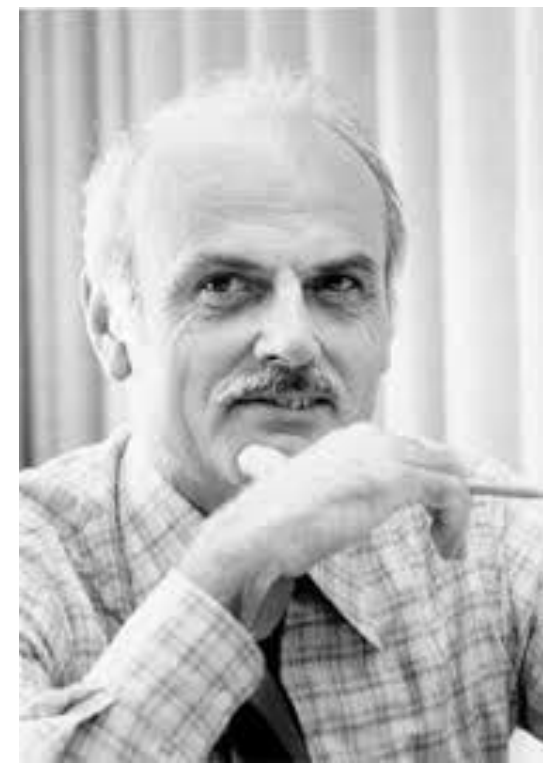


NORMALIZAÇÃO

Disciplina: Banco de Dados

NORMALIZAÇÃO

- Conceito introduzido por Edgar Frank Codd em 1970.
- Utiliza um processo matemático formal baseado na teoria dos conjuntos.
- Aplica um conjunto de regras sobre as tabelas de um banco de dados para verificar se estas foram corretamente projetadas.



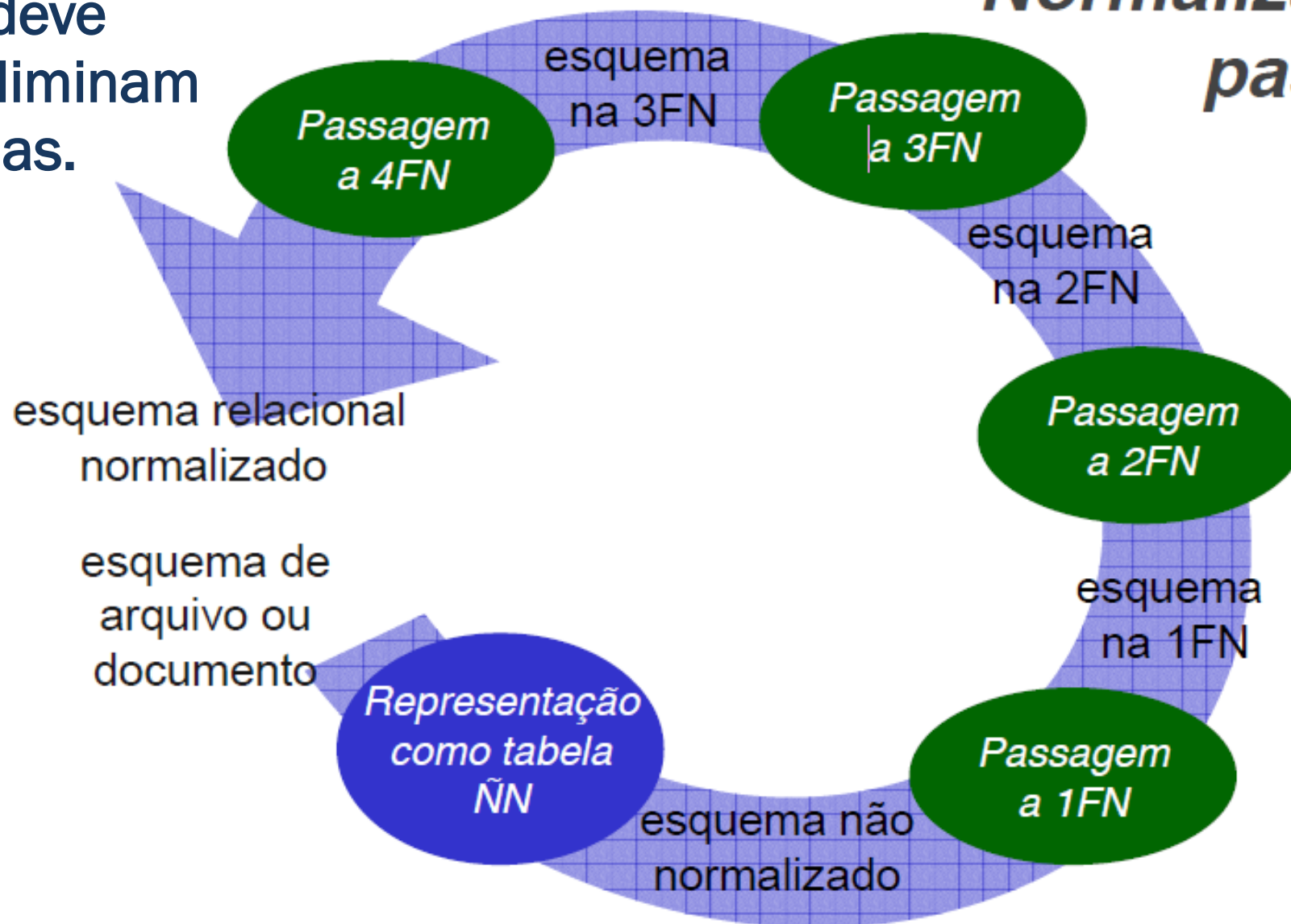
NORMALIZAÇÃO

- **Objetivo principal:** é um conjunto de regras que visa melhorar a qualidade do banco de dados:
 - *reduzir a redundância;*
 - *aumentar a integridade de dados;*
 - *obter tabelas mais evolutíveis e consistentes;*
 - *otimizar o desempenho do banco de dados.*

- Examina os atributos de uma entidade, com o objetivo de evitar erros que podem ser gerados nas atividades de inclusão, alteração e exclusão no banco de dados.

Uma forma normal é um conjunto de regras que uma tabela deve obedecer, que eliminam redundâncias.

Normalização passos



DICA DE OURO

- No Modelo Relacional de Dados “não devemos misturar assuntos em uma mesma Tabela”
- Exemplo: na tabela CLIENTES devemos colocar somente campos relacionados com o assunto Clientes. Não devemos misturar campos relacionados com outros assuntos, tais como Pedidos, Produtos, etc.
- Essa “mistura de assuntos” em uma mesma tabela, acaba por gerar repetições desnecessárias, bem como inconsistência dos dados.

1FN - PRIMEIRA FORMA NORMAL

- Uma entidade está na Primeira Forma Normal se não houver grupo de dados repetidos, ou seja, se todos os valores forem únicos.
- O que significa que a primeira forma normal não permite repetições ou campos que tenha mais de um valor.

Para aplicar a 1FN, devemos:

- Identificar a chave primária da entidade;
- Identificar o grupo repetitivo e removê-lo da entidade;
- Criar uma nova entidade com a chave primária da entidade anterior e o grupo repetitivo.

NÃO NORMALIZADA

1

Código_cliente	Nome	Telefone	Endereço
C001	José	9563-6352 9847-2501	Rua Seis, 85 Morumbi 12536-965
C002	Maria	3265-8596	Rua Onze, 64 Moema 65985-963
C003	Janio	8545-8956 9598-6301	Praça ramos Liberdade 68858-633

Os clientes possuem Rua, CEP e Bairro, e essas informações estão na mesma célula da tabela “Endereço”. Logo a tabela não está na 1FN, porque tem mais de um valor no mesmo campo.

Normalizada na 1FN:

2

Código_cliente	Nome	Telefone	Rua	Bairro	Cep
C001	José	9563-6352 9847-2501	Rua Seis, 85	Morumbi	12536-965
C002	Maria	3265-8596	Rua Onze, 64	Moema	65985-963
C003	Janio	8545-8956 9598-6301	Praça ramos	Liberdade	68858-633

Para normalizar, deveremos colocar cada informação em um campo (coluna) diferente: Rua, Bairro e CEP.

1

NÃO NORMALIZADA

CodProj	Tipo	Descr	CodEmp	Nome	Cat	Sal	DataIni	TempoAl
LSC001	Novo desenv.	Sistema de Estoque	2146	João	A1	4	01-11-2019	24
LSC001	Novo desenv.	Sistema de Estoque	3145	Silvia	A2	6	02-10-2018	24
LSC001	Novo desenv.	Sistema de Estoque	6126	José	B1	9	03-10-2019	18
LSC001	Novo desenv.	Sistema de Estoque	1214	Carlos	A2	6	01-11-2018	18
LSC001	Novo desenv.	Sistema de Estoque	8191	Maria	A1	4	01-11-2019	12
PAG02	Manutenção	Sistema de RH	8191	Maria	A1	4	01-05-2020	12
PAG02	Manutenção	Sistema de RH	4113	João	A2	6	04-01-2018	24
PAG02	Manutenção	Sistema de RH	6126	José	B1	9	01-11-2019	12

Para cada projeto (CodProj) temos mais de um empregado (CodEmp) trabalhando, portanto envolvendo repetições.

Para normalizar, deveremos eliminar as repetições, separando os dados em duas tabelas.

NORMALIZADA na 1FN:

A tabela **Projetos** com a identificação dos projetos e a tabela **Empregados** para identificar os funcionários que atuam em cada projeto.

2

Proj

CodProj	Tipo	Descr
LSC001	Novo desenv.	Sistema de Estoque
PAG02	Manutenção	Sistema de RH

Func

CodEmp	Nome	Cat	Sal
2146	João Santos	A1	4
3145	Silvia	A2	6
6126	José	B1	9
1214	Carlos	A2	6
8191	Maria	A1	4
4113	João Faria	A2	6

ProjEmp

Codigo	CodEmp	DataIni	TempoAI	CodProj
1	2146	01-11-2019	24	LSC001
2	3145	02-10-2018	24	LSC001
3	6126	03-10-2019	18	LSC001
4	1214	01-11-2018	18	LSC001
5	8191	01-11-2019	12	LSC001
6	8191	01-05-2020	12	PAG02
7	4113	04-01-2018	24	PAG02
8	6126	01-11-2019	12	PAG02

2FN - SEGUNDA FORMA NORMAL

Uma tabela está na segunda forma normal quando, além de estar na 1FN, não contém dependências parciais.

A passagem para a 2FN visa eliminar um certo tipo de redundância de dados.

Proj

<u>CodProj</u>	Tipo	<u>Descr</u>
LSC001	Novo <u>desenv.</u>	Sistema de Estoque
LSC001	Manutenção	Sistema de RH

ProjEmp

<u>CodEmp</u>	Nome	<u>Cat</u>	Sal	<u>DataIni</u>	<u>TempoAl</u>	<u>CodProj</u>
2146	João	A1	4	01-11-2019	24	LSC001
3145	Silvia	A2	6	02-10-2018	24	LSC001
6126	José	B1	9	03-10-2019	18	LSC001
1214	Carlos	A2	6	01-11-2018	18	LSC001
8191	Maria	A1	4	01-11-2019	12	LSC001
8191	Maria	A1	4	01-05-2020	12	PAG02
4113	João	A2	6	04-01-2018	24	PAG02
6126	José	B1	9	01-11-2019	12	PAG02

Exemplo:

Na tabela Projeto_Empregado, os dados referentes a empregados (Nome, Cat e Sal) aparecem várias vezes repetidos para os empregados que trabalham em mais de um projeto.

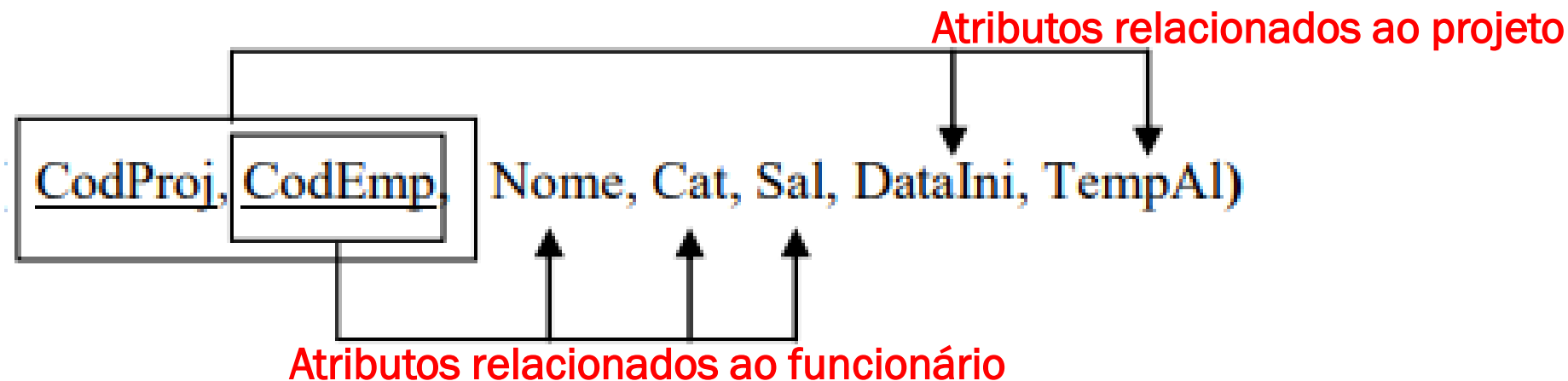
2FN - SEGUNDA FORMA NORMAL

Procedimento:

1. Identificar os atributos que não são dependentes da chave primária.
2. Remover da entidade todos esses atributos identificados e criar uma nova entidade com eles.

2FN - SEGUNDA FORMA NORMAL

- Uma tabela que não se encontre na 2FN possui dependências funcionais parciais, ou seja, **contém colunas não chave que dependem apenas de uma parte da chave primária**.
- Obviamente que se uma tabela está na 1FN e possui apenas uma coluna como chave primária (chave primária simples), então ela não conterá dependências parciais. Sendo assim, ela também já está na 2FN. É o caso da tabela Proj.
- Note as dependências funcionais na tabela ProjEmp.



2FN - SEGUNDA FORMA NORMAL

- Para passar a tabela **ProjEmp** para a 2FN (ou seja, eliminar as dependências parciais) é necessário dividir a mesma em duas tabelas com o seguinte esquema:

ProjEmp(CodProj, CodEmp, DataIni, TempAl)

Emp(CodEmp, Nome, Cat, Sal)

- Assim, o modelo relacional correspondente ao arquivo em questão, na 2FN fica:

Proj(CodProj, Tipo, Descr)

ProjEmp(CodProj, CodEmp, DataIni, TempAl)

Emp(CodEmp, Nome, Cat, Sal)

- O conteúdo das tabelas é exibido a seguir:

Proj:

CodProj	Tipo	Descr
LSC001	Novo Desenv.	Sistema de Estoque
PAG02	Manutenção	Sistema de RH

ProjEmp:

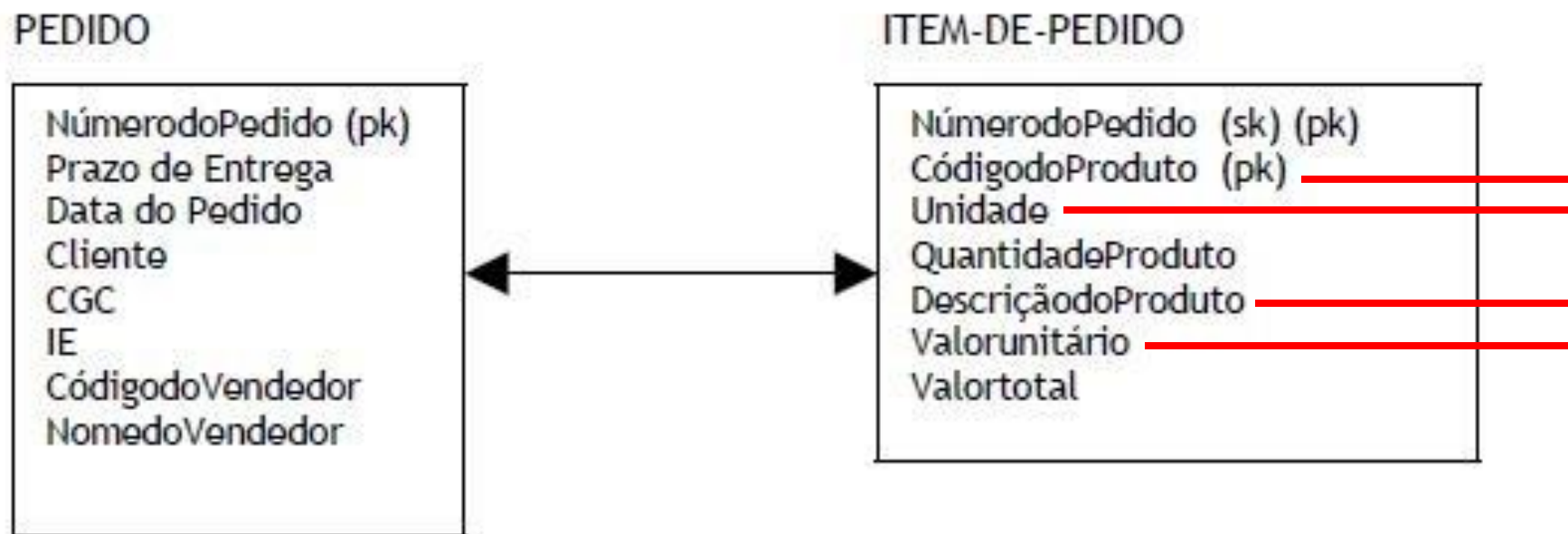
CodProj	CodEmp	DataIni	TempAl
LSC001	2146	01-11-1991	24
LSC001	3145	02-10-1991	24
LSC001	6126	03-10-1992	18
LSC001	1214	01-11-1991	18
LSC001	8191	01-11-1992	12

Emp:

CodEmp	Nome	Cat	Sal
2146	João	A1	4
3145	Sílvio	A2	6
6126	José	B1	9
1214	Carlos	A2	6
8191	Mário	A1	4
4112	João	A2	6

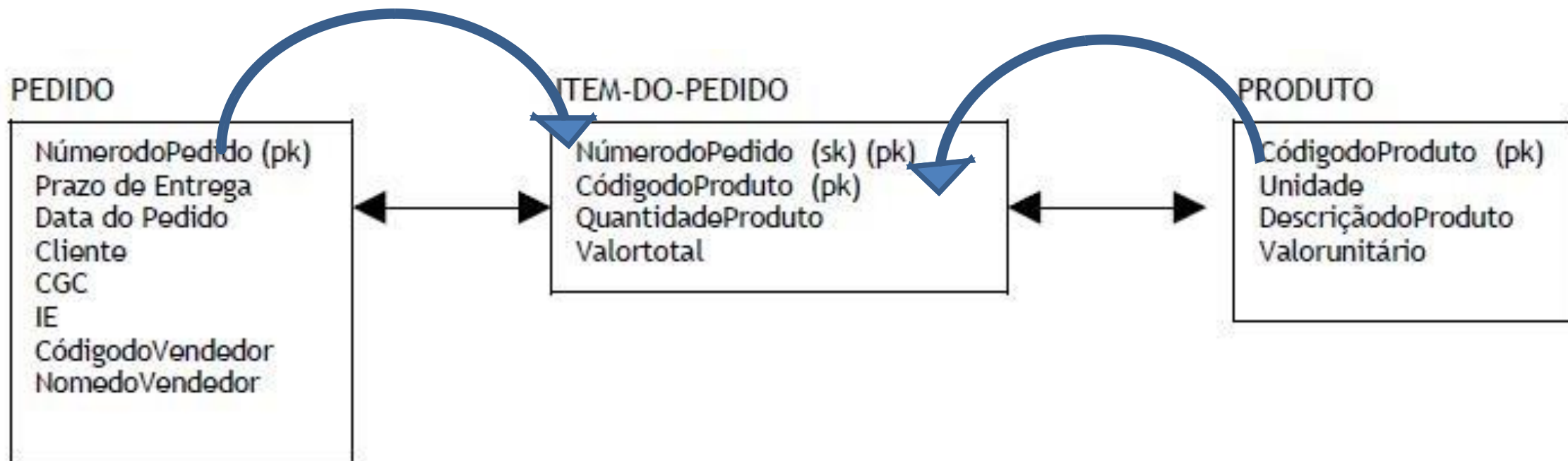
Outro exemplo:

- A entidade ITEM-DE-PEDIDO possui chave primária e notamos que os atributos UNIDADE, DESCRICAO-DO-PRODUTO e VALOR-UNITÁRIO dependem de forma parcial do atributo CODIGO-DOPRODUTO, que faz parte da chave primária.
- Quando aplicamos a 2FN sobre ITEM-DO-PEDIDO, será criada a entidade PRODUTO que herdará os atributos UNIDADE, DESCRICAO-DO-PRODUTO e VALOR-UNITARIO e terá como chave primária CODIGO-DO-PRODUTO.



Outro exemplo:

- Quando aplicamos a 2FN sobre ITEM-DO-PEDIDO, será criada a entidade PRODUTO que herdará os atributos UNIDADE, DESCRICAO-DO-PRODUTO e VALOR-UNITARIO e terá como chave primária CODIGO-DO-PRODUTO.
- Imaginem cadastrar um ITEM-DO-PEDIDO se ainda não foi cadastrado o PRODUTO?



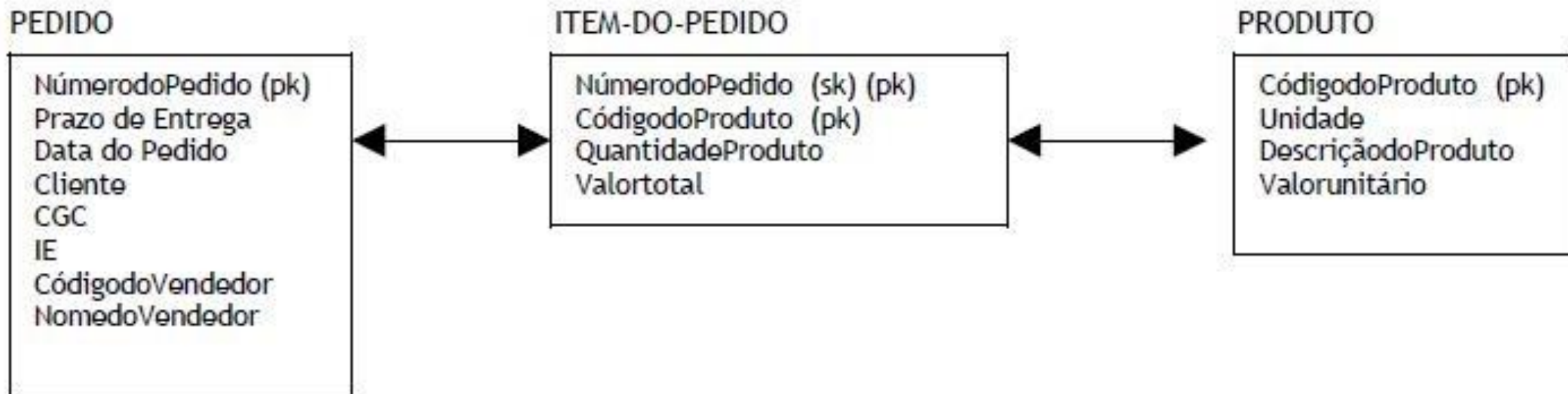
3FN - TERCEIRA FORMA NORMAL

Uma tabela está na terceira forma normal quando, além de estar na 2FN, não contém dependências transitivas, ou seja, nenhuma coluna não chave depender de outra coluna não-chave.

Não deve conter colunas que sejam o resultado de algum cálculo sobre outra coluna.

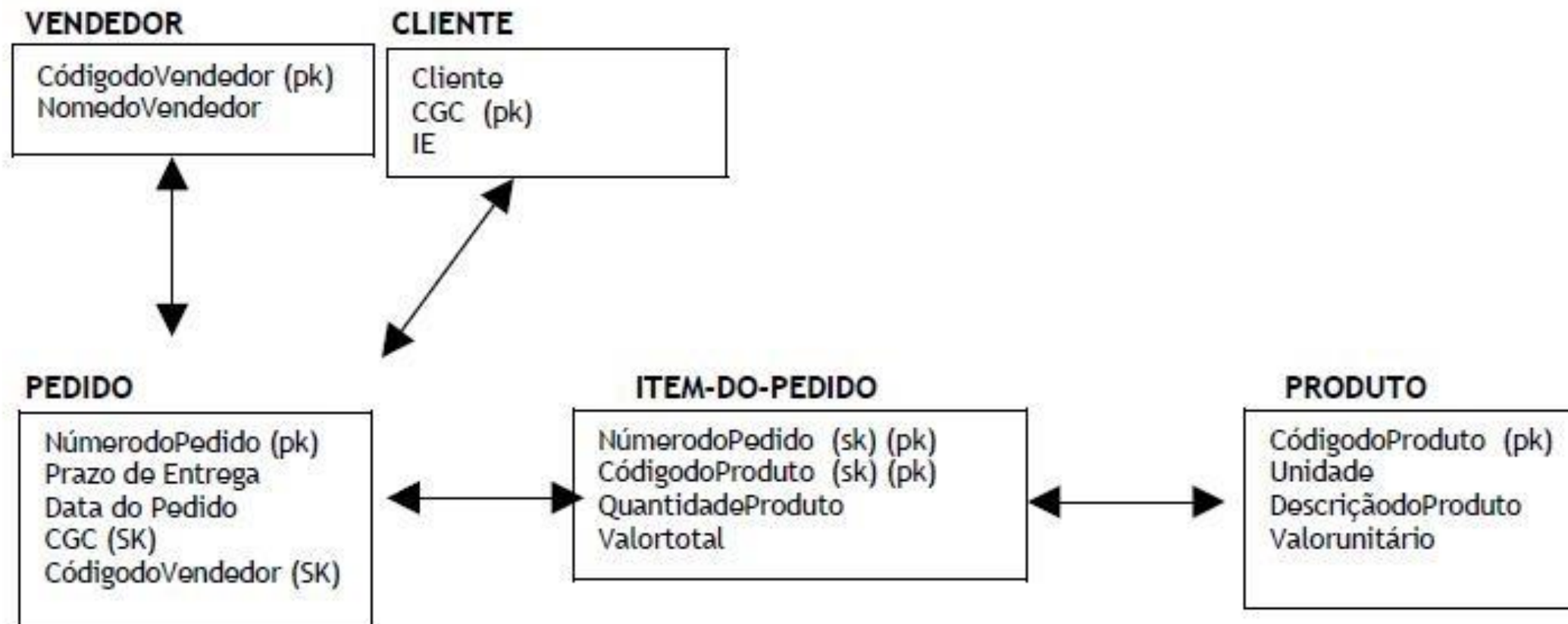
3FN - TERCEIRA FORMA NORMAL

Na tabela PEDIDO, notamos que a coluna Nome do Vendedor depende transitivamente da coluna Código do Vendedor, que não pertence à chave primária Número do Pedido(pk)



3FN - TERCEIRA FORMA NORMAL

Para eliminarmos esta anomalia devemos criar a tabela **VENDEDOR** com seus colunas. Da mesma forma e pelo mesmo motivo, também criaremos a tabela **CLIENTE** com suas colunas, cada tabela com seu assunto específico.



3FN - TERCEIRA FORMA NORMAL

A coluna **Total** é dependente, ou é resultado, da multiplicação das colunas **Preço** e **Quantidade**, ou seja, a coluna total tem dependência transitiva de colunas que não fazem parte da chave primária, ou mesmo candidata da tabela.

Itens do pedido				
Pedido	Item	Preço	Quantidade	Total
15	102	9,25	2	18,5
15	132	1,3	5	6,5

Aqui, após o atributo/coluna *Total* ser excluído da tabela, ela já na 3ª Forma Normal. Esse atributo pode ser movido para outra tabela referenciando a antiga.

Itens do pedido			
Pedido	Item	Preço	Quantidade
15	102	9,25	2
15	132	1,3	5

3FN - TERCEIRA FORMA NORMAL

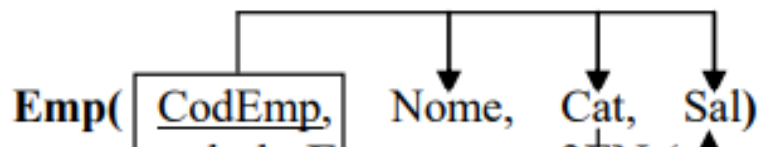
Na terceira forma normal vamos eliminar aqueles campos que podem ser obtidos pela equação de outros campos da mesma tabela.

COD PEDIDO	COD PRODUTO	QUANTIDADE	VALOR UNITÁRIO	VALOR TOTAL
1	A1	10	10,00	100,00
1	A2	15	6,00	90,00
1	A3	5	4,00	20,00
2	A1	3	8,00	30,00
2	A3	2	3,50	8,00
3	A2	2	7,00	12,00
3	A3	4	5,00	16,00
4	A1	6	8,33	50,00
5	A2	6	6,00	36,00

COD PEDIDO	COD PRODUTO	QUANTIDADE	VALOR UNITÁRIO
1	A1	10	10,00
1	A2	15	6,00
1	A3	5	4,00
2	A1	3	8,00
2	A3	2	3,50
3	A2	2	7,00
3	A3	4	5,00
4	A1	6	8,33
5	A2	6	6,00

3FN - TERCEIRA FORMA NORMAL

- A passagem para a 3FN visa eliminar um outro tipo de redundância de dados.
- Para exemplificar, vamos supor que na tabela **Emp** o salário de um empregado (coluna Sal) seja determinado pela sua categoria funcional (coluna Cat).
- Sendo assim, a informação de que um dado salário é pago para uma determinada categoria vai se repetir na tabela tantas vezes quantos forem os empregados que possuírem aquela categoria.



- Para passar a tabela **Emp** para a 3FN (ou seja, eliminar as dependências transitivas) é necessário dividir a mesma em duas tabelas com o seguinte esquema:

Emp(CodEmp, Nome, Cat)
Cat(Cat, Sal)

- Assim, o modelo relacional correspondente ao arquivo em questão, na 3FN fica:

Proj(CodProj, Tipo, Descr)
ProjEmp(CodProj, CodEmp, DataIni, TempAl)
Emp(CodEmp, Nome, Cat)
Cat(Cat, Sal)

3FN - TERCEIRA FORMA NORMAL

Proj:

CodProj	Tipo	Descr
LSC001	Novo Desenv.	Sistema de Estoque
PAG02	Manutenção	Sistema de RH

ProjEmp:

CodProj	CodEmp	DataIni	TempAl
LSC001	2146	01-11-1991	24
LSC001	3145	02-10-1991	24
LSC001	6126	03-10-1992	18
LSC001	1214	01-11-1991	18
LSC001	8191	01-11-1992	12
PAG02	8191	01-05-1993	12
PAG02	4112	04-01-1991	24
PAG02	6126	01-11-1992	12

Emp:

CodEmp	Nome	Cat
2146	João	A1
3145	Sílvio	A2
6126	José	B1
1214	Carlos	A2
8191	Mário	A1
4112	João	A2

Cat:

Cat	Sal
A1	4
A2	6
B1	9

EXEMPLO

Estrutura Não Normalizada	3FN
Nº Cliente	Cliente
Nome	<u>Nº Cliente</u>
Morada	Nome
localidade	Morada
Codigo Postal	Localidade
Empréstimos*	Cod_Postal
Cod_Agencia	
N_emprestimo	Agência
Valor Empréstimo	<u>cod_agencias</u>
Taxa de Juro	Nome Agência
Nome Agência	Nome Gerente
Nome Gerente	
	Empréstimos
	<u>Nº Empréstimos</u>
	Valor Empréstimo
	Taxa de Juro
	<u>Cod_Cliente</u>
	<u>Cod_Agencia</u>