### Banco de Dados I

#### Prof. Dr. Ronaldo Celso Messias Correia

ronaldo.correia@unesp.br





# Dependência Funcional e Normalização

## Qualidade de Projeto

- > O objetivo é evitar os problemas que podem provocar falhas no projeto do banco de dados, bem como eliminar a mistura de assuntos e as correspondentes redundâncias desnecessárias de dados
- Erros encontrados:
  - Repetição de dados
  - Perda de dados
- Como avaliar se um esquema de relação é bom ou ruim?

## Qualidade de Projeto

- Medidas informais para mensurar a qualidade de um projeto de esquema de relação:
  - Semântica de atributos
  - Redução de valores redundantes nas tuplas
  - Redução de valores nulos nas tuplas
  - Prevenção de tuplas espúrias ou erradas

#### Semântica de Atributos

- Modelar um esquema de relação de modo que seja fácil explicar seu significado (ou semântica)
- Não combine os atributos de diferentes tipos de entidades e relacionamentos dentro de uma única relação
- Se um esquema de relação corresponder a um tipo de entidade ou a um tipo de relacionamento, haverá uma justificativa direta para seu significado
- Problemas de ambiguidade semântica a relação não poderá ser explicada facilmente

#### Semântica de Atributos

EMP\_PROJ (EMPREGADO\_PROJETO)

COD_EMP	ENOME	PNUMERO	HORAS	PNOME	PLOCALIZAÇÃO

➤ EMP\_DEPT (EMPREGADO\_DEPARTAMENTO)

COD_EMP	ENOME	DATANASC	ENDEREÇO	DNUMERO	DNOME	COD_GER

Embora não haja nada logicamente errado com essas duas relações, elas são consideradas pobres

## Redução de valores redundantes

- Agrupar os atributos em esquemas de relações tem um efeito significativo no espaço de armazenamento, mesmo com a repetição de atributos
- Modelar esquemas de relações básicas de forma que nenhuma anomalia de inserção, exclusão ou alteração possa ocorrer na relação
- Se houver a possibilidade de ocorrer alguma anomalia, registre-a claramente e tenha certeza que os programas que atualizam o banco de dados operação corretamente

### **Anomalias** EMP\_PROJ

COD_EMP	ENOME	PNUMERO	HORAS	PNOME	PLOCALIZAÇÃO
1	José	P_01	20	Lixo_eletrônico	Bloco 3
1	José	P_02	10	Evidência	Bloco 2
2	Maria	P_01	15	Lixo_eletrônico	Bloco 3
3	Carlos	P_03	13	FCTable	Bloco 2
1	Daniel	P_01	20	Cidades Inteligentes	

- Atualização:
  - Mudança no nome do projeto requer várias mudanças
- ➤ Inserção:
  - Para inserir um novo empregado é necessário incluir os valores dos atributos para o Projeto
  - É difícil incluir um Projeto que ainda não tenha empregado
- ➤ Exclusão:
  - Ao apagar um empregado, às informações relativas ao projeto serão apagadas

## Redução de valores nulos

- Até onde for possível, evite colocar os atributos em uma relação básica cujos valores frequentemente possam ser nulos
- Se os nulos forem inevitáveis, tenha certeza de que eles se aplicam somente em casos excepcionais e não na maioria das tuplas da relação

## Geração de tuplas ilegítimas

- Projete os esquemas de relações de forma que possam ser unidos (join) com igualdade de condições sobre os atributos que sejam chaves primárias ou chaves estrangeiras, de modo a garantir que nenhuma tupla ilegítima seja gerada
- > Evite as relações que contenham o relacionamento entre os atributos que não sejam combinações (chave estrangeira, chave primária)

## Dependência Funcional

- É o conceito mais importante da teoria de projetos de esquemas relacionais
- É uma restrição entre dois conjuntos de atributos do banco de dados
- Definição: Um atributo B de um esquema de relação R é funcionalmente dependente de outro atributo A de R se um valor de A determina um único valor para B em qualquer momento
- Se B é funcionalmente dependente de A, então A determina funcionalmente B
- ➤ Notação: A → B
- É uma propriedade semântica, identificada pelo projetista da base de dados
- Pode ser verificada na instância da base, mas nunca definida a partir dela

## Dependência Funcional

```
Aluno = {RA, Nome, Idade, Curso} {<201001, Paulo, 25, computação>, <200903, Ana, 18, eletrônica>, <201005, Silvio, 18, odontologia>, <201110, Zenir, 25, computação>, <221310, Paulo, 20, computação>,}
```

Não é interessante ter um atributo Idade!!! PQ?

Solução: Data de Nascimento!!!

- A relação Alunos atende as seguintes DFs?
  - Nome → Curso (não)
  - Nome → Idade (não)
  - Curso → Idade (não)
  - Idade → Curso (não)
  - $\blacksquare$  RA  $\longrightarrow$  Nome (sim)

## Normalização

- O processo de normalização sujeita um esquema de relação a uma série de testes para "certificar-se" de que ele satisfaça uma certa forma normal
  - Forma Normal: definição das relações baseada na análise de dependências funcionais
- A normalização pode ser vista como o processo de análise de determinados esquemas de relações com base em suas Dependências Funcionais e chaves primárias para alcançar as propriedades desejáveis
- A qualidade de um projeto de banco de dados pode ser medida a partir da verificação da **forma normal** que ele alcança

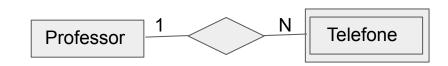
## Normalização

- É um processo de análise e adequação dos esquemas de relações com o objetivo de minimizar redundâncias e anomalias de inserção, exclusão e alteração
- No processo de normalização, esquemas de relação que não satisfazem certas condições, são decompostos em esquemas de relação menores que satisfazem as tais condições e ainda possuem algumas propriedades desejáveis.
  - Os esquemas insatisfatórios são decompostos em relações menores que passam nos testes
- Tipos de formas normais:
  - 1FN, 2FN e 3FN

- Uma relação está na 1FN se:
  - Parte da definição formal do Modelo Relacional
  - Todos os atributos da relação devem ser atômicos e monovalorados
  - Cada ocorrência da chave primária deve corresponder a uma e somente uma informação de cada atributo, ou seja, a entidade não deve conter grupos repetitivos (multivalorados)
  - Gerar uma nova relação contendo o grupo de repetição e a chave primária da relação original
  - Decompor em tantas entidades quanto for o número de conjuntos de atributos repetitivos

Ronaldo Celso Messias Correia - FCT/UNESP

- Atributos não atômicos (Endereço)
  - Alunos =  $\{RA, Nome, Data\_Nasc, Endereço, Disciplinas\}$ 
    - Alunos = {RA, Nome, Data\_Nasc, Rua, Nr, Bairro, Cidade, CEP, UF, Disciplinas}
- Atributos multivalorados
  - Alunos =  $\{RA, Nome, Data\_Nasc, Disciplinas\}$ 
    - Alunos = {RA, Nome, Data\_Nasc}
    - Matricula Disciplinas = {RA, Disciplina}
  - Nas novas entidades criadas, a chave primária é a concatenação da chave primária da entidade original mais o(s) atributo(s) do grupo repetitivo visualizado com a chave primária deste grupo
  - Professor =  $\{\underline{ID}, \text{Nome, e-mail, telefones}\}$ 
    - Professor =  $\{\underline{ID}, \text{Nome, e-mail}\}$
    - Professor\_telefone= {<u>ID</u>, <u>Telefone</u>}



```
Alunos = {RA, Nome, Data_Nasc, Disciplinas}
(1010, Jose da Silva, 03/01/1988, MAT-1120,
2320, Maria Conceição, 13/10/2005, MAT-1120,
1010, Jose da Silva, 03/01/1998, MAT-2110,
1010, Jose da Silva, 03/01/1998, MAT-3130,
2320, Maria Conceição, 13/10/2005, MAT-2120)
```

```
Alunos = {<u>RA</u>, Nome, Data_Nasc}
(1010, Jose da Silva, 03/01/1998,
2320, Maria Conceição, 13/10/2005)
```

```
Matricula_Disciplina (<u>RA, Disciplina</u>)
(1010, MAT-1120,
1010, MAT-2110,
1010, MAT-3130,
2320, MAT-1120,
2320, MAT-2120)
```

- $\rightarrow$  Professor = {<u>ID</u>, Nome, e-mail, telefone1, telefone2, telefone3}
  - Não tem a quantidade exata de telefones por professores
  - Um professor pode ter somente um telefone (atributos nulos)

OU

Um professor pode ter 5 telefones mas não tem como inserir

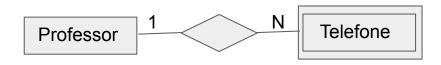
Professor

Telefone

Interessante somente quando souber a quantidade exata de telefones

- Professor = {<u>ID</u>, Nome, e-mails, telefones}
  - Professor = {<u>ID</u>, Nome, e-mail1, e-mail2, e-mail3, ....., e-mailn}
    - (01, Ronaldo Correia, <u>ronaldo.correia@unesp.br</u>, <u>ronaldocmcorreia@gmail.com</u>, null, ...., null;
    - (02, Murilo Correia, <u>murilofcorreia@gmail.com</u>, null, null, .... null)

- Professor\_email(<u>ID, e-mail</u>)
  - (01, ronaldo.correia@unesp.br;
  - (01, ronaldocmcorreia@gmail.com,
  - (02, murilofcorreia@gmail.com)



Cliente 1 Pedido

Pedido (<u>cod\_cliente</u>, nome\_cliente, telefone1, telefone2, endereco, cod\_produto, nome\_produto, preco, quantidade)
Produto

Cod_cliente	Nome_client	Telefone1	Telefone2	Endereço	Cod_produto	Nome_produto	Preço	Quant
1	Murilo	99101001	99101002	Rua A	5412	Arroz	5,00	2
1	Murilo	99101001	99101002	Rua A	4312	<del>Feijão</del>	9,00	1
2	Elaine	98712325	98712327	Rua B	4312	Feijão	9,00	2
3	Pedro	99544123		Rua C	3212	Sardinha	6,50	1
3	Pedro	99544123		Rua C	5412	Arroz	5,00	1

- Pedido(<u>cod\_cliente</u>, nome\_cliente, telefone1, telefone2, Rua, Nr, Bairro, Cidade, CEP, UF)
- Itens\_Pedido(<u>cod\_cliente, cod\_produto</u>, nome\_produto, preco, quantidade)

- Uma relação está na 2FN se:
  - Está na 1FN
  - Não existe atributo não chave que é dependente de somente uma parte da chave primária
  - Dependência funcional Total
- Dependência Funcional Total X Y
  - Na ocorrência de uma chave primária concatenada, dizemos que um atributo ou um conjunto de atributos depende de forma completa ou total desta chave primária concatenada, se e somente se, a cada valor da chave (e não parte dela) está associado um valor para cada atributo
  - Se a remoção de qualquer atributo A de X implica que a dependência não mais será assegurada

Ronaldo Celso Messias Correia - FCT/UNESP

A resolução da aplicação da segunda forma normal é realizada através da exclusão dos atributos que não dependem totalmente da chave primária, da tabela original, e constituindo-se com estes uma nova tabela, que terá como chave primária o atributo participante da chave primária da tabela origem

Ronaldo Celso Messias Correia - FCT/UNESP

- Cliente(<u>cod\_cliente</u>, nome\_cliente, telefone1, telefone2, Rua, Nr, Bairro, Cidade, CEP, UF)
  - Não possui chave primária composta
- Pedido(cod\_cliente, cod\_produto, nome\_produto, preco, preco\_venda, quantidade)
  - cod\_produto → nome\_produto, preço (dependência parcial)
  - cod\_cliente, cod\_produto → quantidade, peco\_venda (dependência total)
- > 2FN:
  - Cliente(<u>cod\_cliente</u>, nome\_cliente, telefone1, telefone2, Rua, Nr, Bairro, Cidade,
     CEP, UF)
  - Produto (<u>cod\_produto</u>, nome\_produto, preco)
  - Pedido (cod\_cliente, cod\_produto, quantidade, preco\_venda) ?

Turma = {Número, Sigla, Sala, NomeDisc, Carga\_horária}
sigla da

disciplina

Depende de parte da chave primária

- Turma = {<u>Número, Sigla,</u> Sala}
- Disciplina = {Sigla, NomeDisc, Carga\_horária}

- Uma relação está na 3FN se:
  - Está na 2FN
  - Não existem atributos não chave que sejam dependentes de outros atributos não chave
  - Dependência transitiva
- Dependência Transitiva
  - Quando um atributo ou conjunto de atributos A depende de outro atributo B que não pertence à chave primária, mas é dependente funcional desta, dizemos que A é dependente transitivo de B

- Funcionario = {Codigo, nome\_func, Coddepto, NomeDepartamento}
   {1, josé da silva, d1, depto de física,
   2, maria, d2, geografia,
   3, carlos, d1, matemática}
- ightharpoonup Funcionario = {Codigo, nome\_func, Coddepto(fk)}
- Departamento = {Coddepto, NomeDepartamento} Funcionário(1, jose da silva, d1; 2, maria, d2; 3; carlos, d1) Departamento (d1, depto de física; d2, geografia)

Disciplina = {Sigla, Nome\_disc, LivroTexto, CodDepto(CodChefeDepto)

- Disciplina = {Sigla, Nome\_disc, LivroTexto, codDepto}
- Chefia= {Depto, CodChefeDepto}

## Definição Geral

- Segunda Forma Normal
  - Um esquema de relação R está na 2FN se cada atributo não primário de R não for parcialmente dependente de nenhuma chave de R
- Terceira Forma Normal
  - Um esquema de relação R está na 3FN se para cada dependência funcional X A, X é uma superchave de R ou A é um atributo primário de R

1) Uma tabela para estar na primeira forma normal (1FN) deve garantir que:

- A. todos os atributos não chave dependam diretamente da chave primária.
- (B) seus atributos sejam atômicos.
- C. nenhuma coluna (atributo) não chave dependa de outra coluna não chave.
- D. todos atributos não chave não dependam diretamente da chave primária.
- E. seus atributos sejam multivalorados.

2) Faça a normalização (1º, 2º e 3º forma normal) dos documentos abaixo e explique os conceitos aplicados em cada forma normal. Definir a chave primária da relação inicial.

Código do Vendedor: 1791

Nome do Vendedor: Anibal da Silva

Prazo de Entrega: 20 dias

Num. Ped.	Cliente	Endereço Cidade/UF	CGC	IE	Cód. Prod.	Unid.	Quant.	Descrição	Val. Unit.	Tot. Prod.	Tot. do Ped.
3445	TCA	R.Meira	11111111	1111111	45	L	20	álcool	5,00	100,00	1799,00
3445	TCA	R.Meira	11111111	1111111	130	М	2	tecido	20,00		1799,00
3445	TCA	R.Meira	11111111	1111111	35	Kg	30	farinha	1,00	30,00	
3445	TCA	R.Meira	11111111	1111111	78	Kg	50	cimento	30,00		
3445	TCA	R.Meira	11111111	1111111	90	L	40	cola	3,00		
3445	TCA	R.Meira	11111111	1111111	39	Kg	3	chumbo	3.00		
2610	Lopes	R. 127	23232323	343434	45	L	50	álcool	5,00	250,00	2650,00
2610	Lopes	R. 127	23232323	343434	78	Kg		cimento	30,00	1410,00	2650,00
2610	Lopes	R. 127	23232323	343434	21	Kg		pregos	5,00	100,00	2650,00
2610	Lopes	R. 127	23232323	343434	98	L		tinta azul	25,00	375,00	2650,00
2610	Lopes	R. 127	23232323	343434	90	L		cola	3,00	45,00	2650,00
2610	Lopes	R. 127	23232323	343434	43	М	10	arame	3,00	30,00	2650,00
2610	Lopes	R. 127	23232323	343434	25	F	10	algodão	2,00	20,00	2650,00
2610	Lopes	R. 127	23232323	343434	65	L	-	querosene	8,00	40,00	2650,00
2610	Lopes	R. 127	23232323	343434	51	М		fio elétrico	13,00	260,00	2650,00
2610	Lopes	R. 127	23232323	343434	74	М		linha 10	4,00	120,00	2650.00

Resolução 2)

Pedido (<u>Cod\_vendedor</u>, Nome\_vendedor, prazo\_entrega, <u>num\_ped</u>, cod\_cliente, nome\_cliente, endereco, CGC, IE, {cod\_prod, Unit, quant, descricao, valor\_unit, total\_prod}, total\_pedido)

Obs: Cuidado - A definição da chave primária da relação altera o resultado o processo de normalização

Resolução: Aplicar a 1FN, 2FN e 3FN

```
Resolução 2)
```

Pedido (<u>Cod\_vendedor</u>, Nome\_vendedor, prazo\_entrega, <u>num\_ped</u>, cod\_cliente, nome\_cliente, endereco, CGC, IE, cod\_prod, Unid, quant, descricao, valor\_unit, total\_prod, total\_pedido)

Endereço - não atômico

Multivalorados - cod prod, Unid, quant, descricao, valor unit, total prod)

1FN: (Atributos atômicos e monovalorados)

Pedido (<u>Cod\_vendedor,</u> Nome\_vendedor, prazo\_entrega, <u>num\_ped,</u> cod\_cliente, nome\_cliente, Rua, NR, Bairro, Cidade, UF, CGC, Inscrição Estadual - IE, total\_pedido)

Itens\_pedido(<u>Cod\_vendedor(pk)</u>, <u>Num\_ped(pk)</u>, <u>cod\_prod(pk)</u>, Unit, quand, descricao, valor\_unit,

total\_prod)

```
Resolução 2)
```

```
Pedido (<u>Cod_vendedor</u>, Nome_vendedor, prazo_entrega, <u>num_ped</u>, cod_cliente, nome_cliente, Rua, NR, Bairro, Cidade, UF, CGC, Inscrição Estadual - IE, total_pedido)
```

Itens\_pedido(<u>Cod\_vendedor (pk)</u>, <u>Num\_ped (pk)</u>, <u>cod\_prod (pk)</u>, Unid, quant, descricao, valor\_unit, total\_prod)

2FN: (atributo não chave depende de parte da chave - chave primária composta)

Pedido (<u>cod\_vendedor</u>, prazo\_entrega, <u>num\_ped</u>, cod\_cliente, nome\_cliente, Rua, NR, Bairro, Cidade, UF, CGC, Inscrição Estadual - IE, total\_pedido)

Vendedor (cod vendedor, nome vendedor

Itens\_pedido (<u>Cod\_vendedor (pk), Num\_ped (pk), cod\_prod (pk)</u>, quant, total\_prod)

Produto (cod prod, descricao, unid, valor\_unit)

#### Resolução 2)

Pedido (<u>cod\_vendedor</u>, prazo\_entrega, <u>num\_ped</u>, cod\_cliente, nome\_cliente, Rua, NR, Bairro, Cidade, UF, CGC, Inscrição Estadual - IE, total\_pedido)

Vendedor (<a href="mailto:cod\_vendedor">cod\_vendedor</a>, nome\_vendedor

Itens\_pedido (<u>Cod\_vendedor (pk)</u>, <u>Num\_ped (pk)</u>, <u>cod\_prod (pk)</u>, quant, total\_prod)

Produto (cod prod, descricao, unid, valor unit)

3FN: atributos não chave que dependem de outros atributos não chave

Pedido (cod\_vendedor, prazo\_entrega, num\_ped, cod\_cliente, total\_pedido)

Cliente (cod\_cliente, nome\_cliente, Rua, NR, Bairro, Cidade, UF, CGC, Inscrição Estadual - IE,

Vendedor (cod vendedor, nome vendedor

Itens\_pedido (Cod\_vendedor(pk), Num\_ped(pk), cod\_prod(pk), quant, total\_prod)

Produto (cod\_prod, descricao, unid, valor\_unit)

3) Faça a normalização (1º, 2º e 3º forma normal) dos documentos abaixo e explique os conceitos aplicados em cada forma normal. Chave primária Número da Nota Fiscal.

COOPERATIVA QUALQUER SERVIÇOS				A FISCAL DE SERV			
CLIENTE ENDERE	ço						
DATA DE	EMISSÃO	/	/	VALOR ISS			
cóprco	ÓDIGO QUANT. DESCRIÇÃO		Processo 7 o		PREÇO		
CODIGO			RIÇAO	UNIT.	TOTAL		
	$\vdash$						
O ISS I	Á ESTÁ INCI	LIÍDO NO I	PRECO	DOS SERVICOS	TOTAL	-	

```
Resolução 3)
```

Nota\_Fiscal (<u>num\_nota</u>, data\_emissao, data\_limite, nome, endereco, valor\_iss, {cod\_produto, qtde, descricao, preco\_unit, total}, total\_geral)

Chave primária: num\_nota

Resolução: Aplicar a 1FN, 2FN e 3FN

```
Resolução 3)
Nota Fiscal (num nota, data emissao, data limite, nome, endereco, valor iss, {cod produto, gtde,
descricao, preco unit, total, total geral)
1FN:
Nota fiscal (<u>num nota</u>, data emissao, data limite, nome, rua, nr, bairro,
cidade, uf, valor iss, total geral)
Itens_nota (<u>num_nota</u>, <u>cod_produto</u>, qtde, descricao, unit, total)
```

Ronaldo Celso Messias Correia - FCT/UNESP

```
Resolução 3)
```

Nota\_fiscal (<u>num\_nota</u>, data\_emissao, data\_limite, nome, rua, nr, bairro, cidade, uf, valor\_iss, total\_geral)

Itens\_nota (<u>num\_nota</u>, <u>cod\_produto</u>, qtde, descricao, preco\_unit, total)

#### 2FN:

Nota\_fiscal (<u>num\_nota</u>, data\_emissao, data\_limite, nome, rua, nr, bairro, cidade, uf, valor\_iss, total\_geral)

```
itens_nota(<u>num_nota</u>, <u>cod_produto</u>, qtde, total) - (<u>total / qtde = preco_unit</u>)

ou

itens_nota(<u>num_nota</u>, <u>cod_produto</u>, qtde, preco_unit) - (<u>preco_unit</u> * qtde = total)
```

Produto(<u>cod\_produto</u>, descricao, preco\_unit)

```
Resolução 3)
Nota fiscal (num nota, data emissao, data limite, nome, rua, nr, bairro, cidade, uf, valor iss,
total geral)
itens_nota(<u>num_nota</u>, <u>cod_produto</u>, qtde, preco_unit) - (<u>preco_unit * qtde = total</u>)
Produto(<u>cod produto</u>, descricao, preco unit)
3EN.
Nota_fiscal ( <u>num_nota_</u>, data_emissao, data_limite, cod_cliente (FK), valor_iss, total_geral)
Cliente ( cod cliente , nome, rua, nr, bairro, cidade, uf)
itens_nota( <u>num_nota</u>, <u>cod_produto</u>, qtde, preco_unit)
Produto(cod produto, descricao, preco unit)
```

Ronaldo Celso Messias Correia - FCT/UNESP

4) Considerando o Modelo proposto:

```
Movimentacao (Nro_movimentacao, Nro-Agencia, Nome_agencia, Endereco_agencia, Gerente_agencia, Nro-Conta, Nome_cliente, CPF, RG_cliente, Fone1_cliente, Fone2_cliente, Endereco_cliente, EstadoCivil_cliente, Cr-Db, Tipo, Data, Hora, Valor, Saldo conta)
```

Aplicar as três primeiras Formas Normais

5) Faça a normalização (1ª, 2ª e 3ª forma normal) dos documentos abaixo e explique os conceitos aplicados em cada forma normal. Chave primária do Cupom Fiscal é GNF.

Cupom\_Fiscal (Data, hora, <u>GNF</u>, COD, {Codigo, Descricao, quantidade, unitario, valor}, qtde\_itens, operadora, NF, PDV, total)

Resolução: A partir da relação aplicar a processo de normalização

ELDORAC AV. MANDEL P	00 S/A GOULART, 2400
PRESIDENTE PRUDENTE -	SP TEL: (018)39022800
CNPJ:62.545.579/0025-00 IM:45583	IE:562091565112

10/06/2007 20:18	GNF:001032 CDD:002	125
CUPOM ITM CODIGO DESCRIÇÃO	FISCAL	
QUANTIDADE UNITARI 1 000923 71PAO FRANC		
00000,145 x	4.00 I 50	/Kg ,58
2 002943 73LEITE LIDER 3 002943 73LEITE LIDER	00001x 1,67 Tg 31	,67
4 045230 08BISC RECH C	00001x 1.09 Te \$1	.67
5 045230 08BISC RECH C		,09

TOTAL	R\$	William Co.	6.10
Dinheiro	www. with the file of the file of	P\$ laurie	10.10
Troco		R\$	4.00
TA=05,60%	T <sub>B</sub> =07.00%	T <sub>C</sub> =08,80%	Tp=12,00%
T=18,00%	T <sub>F</sub> =25,00%	Se=05.00%	.0-12/00%

NR PDV.:14 DRT: 20134