



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

Álgebra Relacional

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)

Álgebra Relacional

- Conjunto de operações teóricas para manipular bancos de dados relacionais
 - **Entrada:** uma ou mais relações
 - **Saída:** uma relação

Porque estudar Álgebra Relacional?

- Os conceitos da álgebra relacional são utilizados em linguagens de manipulação de dados (ex: SQL)
 - Base para otimizar e implementar as consultas no banco
 - Formalização das operações realizadas em bancos de dados relacionais

Classificação das Operações da Álgebra Relacional

- As operações da álgebra relacional podem ser classificadas como:
 - **Operações específicas de Bancos de Dados Relacionais**
 - Seleção, Projeção, Junção
 - **Operações da Teoria dos Conjuntos da Matemática**
 - União, Intersecção, Subtração, Produto cartesiano

Operação Seleção (σ)

- Escolhe um subconjunto de registros (tuplas) de uma relação **R** que satisfaz uma **condição de seleção**



Operação Seleção (σ)

- Escolhe um subconjunto de registros (tuplas) de uma relação **R** que satisfaz uma **condição de seleção**



ATENÇÃO:

A saída da operação seleção (σ) é uma relação com os MESMOS atributos de **R**.
Isto é, a saída tem o mesmo **grau de relação** que R.

Exemplo de Operação Seleção (σ)

Funcionario_Departamento				
cpf_func	nome_func	data_nasc	salario	cpf_gerente
777.888.999-00	Carlos	02-05-1985	3100	504.231.854-78
111.222.333-44	Jussara	05-07-1995	2000	854.632.140-98
000.555.777-11	Maria	20-02-1970	4000	784.201.632-11

- $\sigma_{\text{salario} > 3000}(\text{Funcionario_Departamento})$

Exemplo de Operação Seleção (σ)

Funcionario_Departamento				
cpf_func	nome_func	data_nasc	salario	cpf_gerente
777.888.999-00	Carlos	02-05-1985	3100	504.231.854-78
111.222.333-44	Jussara	05-07-1995	2000	854.632.140-98
000.555.777-11	Maria	20-02-1970	4000	784.201.632-11

- $\sigma_{\text{salario} > 3000}(\text{Funcionario_Departamento})$

Exemplo de Operação Seleção (σ)

Funcionario_Departamento				
cpf_func	nome_func	data_nasc	salario	cpf_gerente
777.888.999-00	Carlos	02-05-1985	3100	504.231.854-78
111.222.333-44	Jussara	05-07-1995	2000	854.632.140-98
000.555.777-11	Maria	20-02-1970	4000	784.201.632-11

- $\sigma_{\text{salario} > 3000}(\text{Funcionario_Departamento})$

cpf_func	nome_func	data_nasc	salario	cpf_gerente
777.888.999-00	Carlos	02-05-1985	3100	504.231.854-78
000.555.777-11	Maria	20-02-1970	4000	784.201.632-11

Exemplo de Operação Seleção (σ)

Funcionario_Departamento				
cpf_func	nome_func	data_nasc	salario	cpf_gerente
777.888.999-00	Carlos	02-05-1985	3100	504.231.854-78
111.222.333-44	Jussara	05-07-1995	2000	854.632.140-98
000.555.777-11	Maria	20-02-1970	4000	784.201.632-11

- $\sigma_{(\text{salario} > 3000 \text{ and } \text{data_nasc} < 01-01-1980)}$ (Funcionario_Departamento)

Exemplo de Operação Seleção (σ)

Funcionario_Departamento				
cpf_func	nome_func	data_nasc	salario	cpf_gerente
777.888.999-00	Carlos	02-05-1985	3100	504.231.854-78
111.222.333-44	Jussara	05-07-1995	2000	854.632.140-98
000.555.777-11	Maria	20-02-1970	4000	784.201.632-11

- $\sigma_{(\text{salario} > 3000 \text{ and } \text{data_nasc} < 01-01-1980)}$ (Funcionario_Departamento)

Exemplo de Operação Seleção (σ)

Funcionario_Departamento				
cpf_func	nome_func	data_nasc	salario	cpf_gerente
777.888.999-00	Carlos	02-05-1985	3100	504.231.854-78
111.222.333-44	Jussara	05-07-1995	2000	854.632.140-98
000.555.777-11	Maria	20-02-1970	4000	784.201.632-11

- $\sigma_{(\text{salario} > 3000 \text{ and } \text{data_nasc} < 01-01-1980)}$ (Funcionario_Departamento)

Exemplo de Operação Seleção (σ)

Funcionario_Departamento				
cpf_func	nome_func	data_nasc	salario	cpf_gerente
777.888.999-00	Carlos	02-05-1985	3100	504.231.854-78
111.222.333-44	Jussara	05-07-1995	2000	854.632.140-98
000.555.777-11	Maria	20-02-1970	4000	784.201.632-11

- $\sigma_{(\text{salario} > 3000 \text{ and } \text{data_nasc} < 01-01-1980)}$ (Funcionario_Departamento)

cpf_func	nome_func	data_nasc	salario	cpf_gerente
000.555.777-11	Maria	20-02-1970	4000	784.201.632-11

Operação Seleção (σ)

- Todos os operadores de comparação podem ser utilizados
 - $>, <, >=, <=, =$
- Número de tuplas do resultado da operação seleção é sempre menor ou igual ao número de tuplas de R
 - $|\sigma_c(R)| \leq |R|$

Operação Seleção (σ)

- Operação seleção é **comutativa**
 - $\sigma_{\text{cond1}}(\sigma_{\text{cond2}}(R)) = \sigma_{\text{cond2}}(\sigma_{\text{cond1}}(R))$
 - *“A ordem das operações de seleção não muda o resultado”*

Exemplo de Comutatividade da Operação Seleção (σ)

Funcionario_Departamento				
cpf_func	nome_func	data_nasc	salario	cpf_gerente
777.888.999-00	Carlos	02-05-1985	3100	504.231.854-78
111.222.333-44	Jussara	05-07-1995	2000	854.632.140-98
000.555.777-11	Maria	20-02-1970	4000	784.201.632-11

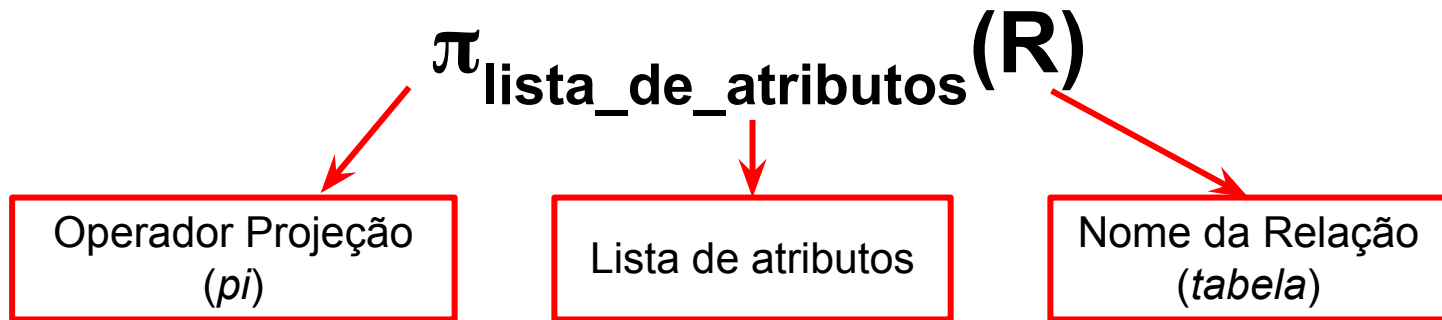
• $\sigma_{\text{salario} > 3000}(\sigma_{\text{data_nasc} < 01-01-1980}(\text{Funcionario_Departamento}))$

$\sigma_{\text{data_nasc} < 01-01-1980}(\sigma_{\text{salario} > 3000}(\text{Funcionario_Departamento}))$

cpf_func	nome_func	data_nasc	salario	cpf_gerente
000.555.777-11	Maria	20-02-1970	4000	784.201.632-11

Operação Projeção (π)

- Considere que as relações são tabelas
 - **Operação Seleção:** escolhe as **linhas** da tabela
 - **Operação Projeção:** escolhe as **colunas** da tabela
- **Operação Projeção** escolhe os atributos de uma relação



Exemplo de Operador Projeção (π)

Funcionario_Departamento				
cpf_func	nome_func	data_nasc	salario	cpf_gerente
777.888.999-00	Carlos	02-05-1985	3100	504.231.854-78
111.222.333-44	Jussara	05-07-1995	2000	854.632.140-98
000.555.777-11	Maria	20-02-1970	4000	784.201.632-11

- $\pi_{\text{cpf_func, nome_func, data_nasc}}(\text{Funcionario_Departamento})$

cpf_func	nome_func	data_nasc
777.888.999-00	Carlos	02-05-1985
111.222.333-44	Jussara	05-07-1995
000.555.777-11	Maria	20-02-1970

Operação Projeção (π)

- O resultado da operação projeção contém **apenas** os atributos especificados na lista de atributos, na mesma ordem
- A operação projeção remove quaisquer tuplas duplicadas
 - Resultado sempre é um conjunto de tuplas distintas
 - O número de tuplas resultantes da projeção é sempre menor ou igual ao número de tuplas em **R**

Exemplo de Operador Projeção (π)

Funcionario_Departamento				
cpf_func	nome_func	data_nasc	salario	cpf_gerente
777.888.999-00	Carlos	02-05-1985	3100	504.231.854-78
111.222.333-44	Maria	05-07-1995	4000	854.632.140-98
000.555.777-11	Maria	20-02-1970	4000	784.201.632-11

- $\pi_{\text{nome_func, salario}}(\text{Funcionario_Departamento})$

Exemplo de Operador Projeção (π)

Funcionario_Departamento				
cpf_func	nome_func	data_nasc	salario	cpf_gerente
777.888.999-00	Carlos	02-05-1985	3100	504.231.854-78
111.222.333-44	Maria	05-07-1995	4000	854.632.140-98
000.555.777-11	Maria	20-02-1970	4000	784.201.632-11

- $\pi_{\text{nome_func, salario}}(\text{Funcionario_Departamento})$

nome_func	salario
Carlos	3100
Maria	4000

Exemplo de Operador Projeção (π)

Funcionario_Departamento				
cpf_func	nome_func	data_nasc	salario	cpf_gerente
777.888.999-00	Carlos	02-05-1985	3100	504.231.854-78
111.222.333-44	Maria	05-07-1995	4000	854.632.140-98
000.555.777-11	Maria	20-02-1970	4000	784.201.632-11

- $\pi_{\text{nome_func}, \text{salario}}(\text{Funcionario_Departamento})$

nome_func	salario
Carlos	3100
Maria	4000

Deveríamos ter duas funcionárias “Maria” com mesmo salário.

Exemplo de Operador Projeção (π)

Funcionario_Departamento				
cpf_func	nome_func	data_nasc	salario	cpf_gerente
777.888.999-00	Carlos	02-05-1985	3100	504.231.854-78
111.222.333-44	Maria	05-07-1995	4000	854.632.140-98
000.555.777-11	Maria	20-02-1970	4000	784.201.632-11

- $\pi_{\text{nome_func}, \text{salario}}(\text{Funcionario_Departamento})$

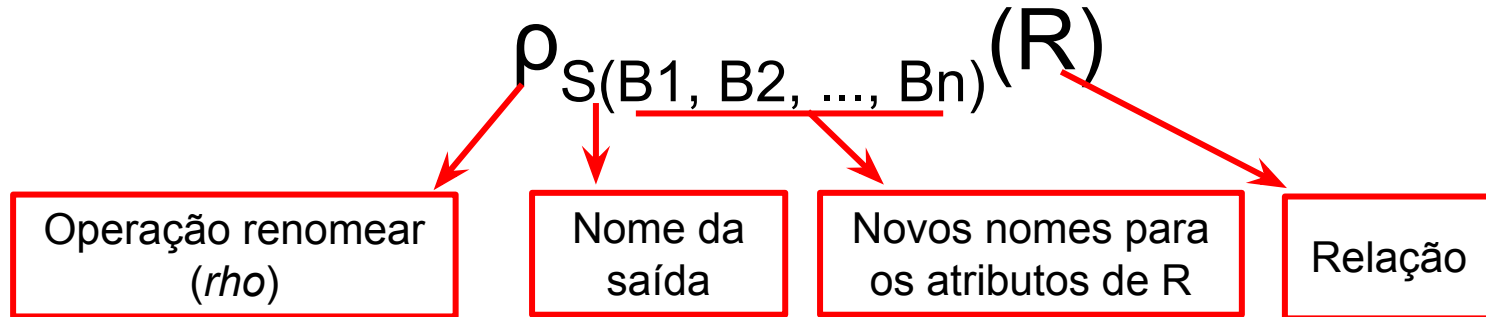
nome_func	salario
Carlos	3100
Maria	4000

Deveríamos ter duas funcionárias “Maria” com mesmo salário.

Porém, a operação projeção removeu as tuplas duplicadas

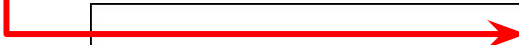
Operação Renomear (ρ)

- Os resultados das operações até agora não tem nome
- É importante nomear os resultados para evitar que tenhamos muitas operações sendo feitas ao mesmo tempo
 - Podemos nomear cada operação



Funcionario_Departamento				
cpf_func	nome_func	data_nasc	salario	cpf_gerente
777.888.999-00	Carlos	02-05-1985	3100	504.231.854-78
111.222.333-44	Jussara	05-07-1995	2000	854.632.140-98
000.555.777-11	Maria	20-02-1970	4000	784.201.632-11

- $\rho_{\text{func_salario}}(\sigma_{\text{salario} > 3000}(\text{Funcionario_Departamento}))$



func_salario				
cpf_func	nome_func	data_nasc	salario	cpf_gerente
777.888.999-00	Carlos	02-05-1985	3100	504.231.854-78
000.555.777-11	Maria	20-02-1970	4000	784.201.632-11

Funcionario_Departamento				
cpf_func	nome_func	data_nasc	salario	cpf_gerente
777.888.999-00	Carlos	02-05-1985	3100	504.231.854-78
111.222.333-44	Jussara	05-07-1995	2000	854.632.140-98
000.555.777-11	Maria	20-02-1970	4000	784.201.632-11

- $\rho_{\text{resultado}}(\pi_{\text{cpf_func, nome_func, data_nasc}}(\text{func_salario}))$

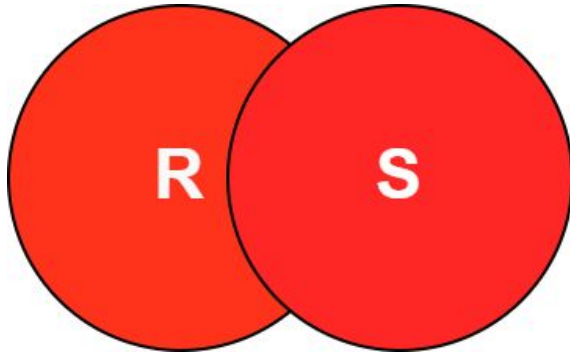
resultado		
cpf_func	nome_func	data_nasc
777.888.999-00	Carlos	02-05-1985
000.555.777-11	Maria	20-02-1970

Exercicio - Seleção, Projeção e Renomear

- Seja uma tabela de Produtos(id, nome, valor, qtd, nome_fornecedor)
- Realize uma consulta usando as operações seleção, projeção e renomear para encontrar os registros que:
 - Possuem **valor > 30** e **qtd < 20**
 - O resultado deve ser nomeado como “reabastecer_produtos” e deve conter apenas os atributos nome, valor e qtd
 - **Dica:** você pode utilizar condições lógicas (AND, OR, NOT)

Operação União (U)

- Inclui todas as tuplas contidas nas relações **R** e **S**
 - Por definição, tuplas duplicadas são removidas



$$Q = R \cup S$$

Funcionario_Departamento				
cpf_func	nome_func	data_nasc	salario	cpf_gerente
777.888.999-00	Carlos	02-05-1985	3100	504.231.854-78
111.222.333-44	Jussara	05-07-1995	2000	854.632.140-98
000.555.777-11	Maria	20-02-1970	4000	784.201.632-11

- $\rho_{\text{func_salario}}(\sigma_{\text{salario} \geq 4000}(\text{Funcionario_Departamento}))$

func_salario				
cpf_func	nome_func	data_nasc	salario	cpf_gerente
000.555.777-11	Maria	20-02-1970	4000	784.201.632-11

Funcionario_Departamento				
cpf_func	nome_func	data_nasc	salario	cpf_gerente
777.888.999-00	Carlos	02-05-1985	3100	504.231.854-78
111.222.333-44	Jussara	05-07-1995	2000	854.632.140-98
000.555.777-11	Maria	20-02-1970	4000	784.201.632-11

- $\rho_{\text{func_nasc}}(\sigma_{\text{data_nasc} > 01-01-1990}(\text{Funcionario_Departamento}))$

func_nasc				
cpf_func	nome_func	data_nasc	salario	cpf_gerente
111.222.333-44	Jussara	05-07-1995	2000	854.632.140-98

func_salario				
cpf_func	nome_func	data_nasc	salario	cpf_gerente
000.555.777-11	Maria	20-02-1970	4000	784.201.632-11

func_nasc				
cpf_func	nome_func	data_nasc	salario	cpf_gerente
111.222.333-44	Jussara	05-07-1995	2000	854.632.140-98

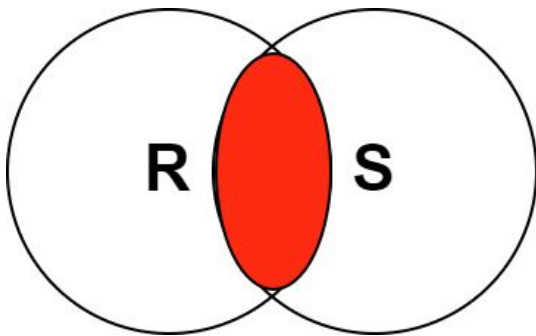
- $\rho_{\text{resultado}}(\text{func_salario} \cup \text{func_nasc})$

Combine as tuplas de
func_salario e *func_nasc*

resultado				
cpf_func	nome_func	data_nasc	salario	cpf_gerente
111.222.333-44	Jussara	05-07-1995	2000	854.632.140-98
000.555.777-11	Maria	20-02-1970	4000	784.201.632-11

Operação Intersecção (\cap)

- Inclui as tuplas que estão em **ambas** as relações **R** e **S**
 - O resultado contém as tuplas que existem em **R** e em **S**



$$Q = R \cap S$$

Funcionario_Departamento			
cpf_func	nome_func	data_nasc	salario
777.888.999-00	Carlos	02-05-1985	3100
111.222.333-44	Jussara	05-07-1995	2000
000.555.777-11	Maria	20-02-1970	4000

Gerente_Departamento			
cpf_func	nome_func	data_nasc	salario
504.231.854-78	José	02-05-1985	7000
000.555.777-11	Maria	20-02-1970	4000
784.201.632-11	Cassandra	10-08-1960	8000

“*Maria*” existe em ambas as relações.
Logo a saída **Q** deve conter a tupla
“*Maria*”

Funcionario_Departamento			
cpf_func	nome_func	data_nasc	salario
777.888.999-00	Carlos	02-05-1985	3100
111.222.333-44	Jussara	05-07-1995	2000
000.555.777-11	Maria	20-02-1970	4000

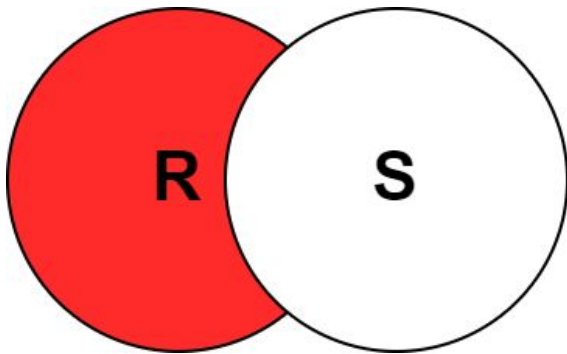
Gerente_Departamento			
cpf_func	nome_func	data_nasc	salario
504.231.854-78	José	02-05-1985	7000
000.555.777-11	Maria	20-02-1970	4000
784.201.632-11	Cassandra	10-08-1960	8000

- $\rho_{\text{resultado}}(\text{Funcionario_Departamento} \cap \text{Gerente_Departamento})$

resultado			
cpf_func	nome_func	data_nasc	salario
000.555.777-11	Maria	20-02-1970	4000

Operação Subtração (-)

- Inclui as tuplas que estão na relação **R** e não estão em **S**
 - O resultado contém as tuplas que existem somente em **R**



$$Q = R - S$$

Funcionario_Departamento			
cpf_func	nome_func	data_nasc	salario
777.888.999-00	Carlos	02-05-1985	3100
111.222.333-44	Jussara	05-07-1995	2000
000.555.777-11	Maria	20-02-1970	4000

Gerente_Departamento			
cpf_func	nome_func	data_nasc	salario
504.231.854-78	José	02-05-1985	7000
000.555.777-11	Maria	20-02-1970	4000
784.201.632-11	Cassandra	10-08-1960	8000

- $\rho_{\text{resultado}}$ (Funcionario_Departamento - Gerente_Departamento)

“Maria” existe em Gerente_Departamento.
Logo a saída Q não deve conter a tupla “Maria”

Funcionario_Departamento			
cpf_func	nome_func	data_nasc	salario
777.888.999-00	Carlos	02-05-1985	3100
111.222.333-44	Jussara	05-07-1995	2000
000.555.777-11	Maria	20-02-1970	4000

Gerente_Departamento			
cpf_func	nome_func	data_nasc	salario
504.231.854-78	José	02-05-1985	7000
000.555.777-11	Maria	20-02-1970	4000
784.201.632-11	Cassandra	10-08-1960	8000

- $\rho_{\text{resultado}}$ (Funcionario_Departamento - Gerente_De

resultado			
cpf_func	nome_func	data_nasc	salario
777.888.999-00	Carlos	02-05-1985	3100
111.222.333-44	Jussara	05-07-1995	2000

Comutatividade e Associatividades de União, Intersecção e Subtração

- As operações união e intersecção são comutativas e associativas
 - $R \cup S = S \cup R$
 - $R \cap S = S \cap R$
 - $(R \cup S) \cup T = R \cup (S \cup T)$
 - $(R \cap S) \cap T = R \cap (S \cap T)$
- A operação subtração é **apenas** associativa
 - $(R - S) - T = R - (S - T)$

Execução de Operações União, Intersecção e Subtração

- As operações união, intersecção e subtração são binárias
 - Cada operação é aplicada a dois conjuntos de tuplas somente
- **Ex:** Sejam as relações R, S, e T:
 - $R \cup S \cup T = (R \cup S) \cup T$
 - $R \cap S \cap T = (R \cap S) \cap T$
 - $R - S - T = (R - S) - T$

Operações União, Intersecção e Subtração

- Uma operação binária é válida se as relações possuírem:
 - **Compatibilidade de União**
 - As duas relações tem o mesmo número de atributos
 - **Compatibilidade de Tipo**
 - Cada par de atributos tem o mesmo domínio

Funcionario_Departamento			
cpf_func	nome_func	data_nasc	salario
777.888.999-00	Carlos	02-05-1985	3100
111.222.333-44	Jussara	05-07-1995	2000
000.555.777-11	Maria	20-02-1970	4000

Gerente_Departamento			
cpf_func	nome_func	data_nasc	salario
504.231.854-78	José	02-05-1985	7000
000.555.777-11	Maria	20-02-1970	4000
784.201.632-11	Cassandra	10-08-1960	8000

Compatibilidade de união:
Ambas as relações tem 4 atributos.

Compatibilidade de tipo:
Cada par de atributos com mesmo nome possui mesmo domínio.

Funcionario_Departamento			
cpf_func	nome_func	data_nasc	salario
777.888.999-00	Carlos	02-05-1985	3100
111.222.333-44	Jussara	05-07-1995	2000
000.555.777-11	Maria	20-02-1970	4000

Compatibilidade de união:
Ambas as relações tem 4 atributos.

Compatibilidade de tipo:
Cada par de atributos com mesmo nome possui mesmo domínio.

Gerente_Departamento			
cpf_func	nome_func	data_nasc	salario
504.231.854-78	José	02-05-1985	7000
000.555.777-11	Maria	20-02-1970	4000
784.201.632-11	Cassandra	10-08-1960	8000

- $\rho_{\text{resultado}}(\text{Funcionario_Departamento} \cap \text{Gerente_Departamento})$

resultado			
cpf_func	nome_func	data_nasc	salario
000.555.777-11	Maria	20-02-1970	4000

Exercício - União, Intersecção, Subtração

- Seja uma tabela de Produtos(id: **int**, nome: **varchar(20)**, valor: **double**, qtd: **int**, id_fornecedor: **int + FK**)
- Realize consultas para encontrar os registros que possuem **valor > 30** e **qtd < 20**
 - Utilize união, intersecção e/ou subtração para construir as consultas
 - Cada consulta pode conter somente uma operação de seleção, e esta NÃO pode ter condições lógicas AND, OR, NOT

Operação Produto Cartesiano (X)

- Também conhecida como operação de:
 - Produto Cruzado
 - Junção Cruzada
- Esta operação combina as tuplas da relação R com as de S

$$Q = R \times S$$

Operação Produto Cartesiano (X)

- Também conhecida como operação de:
 - Produto Cruzado
 - Junção Cruzada
- Esta operação combina as tuplas da relação R com as de S

$$Q = R \times S$$

Não é necessário ter compatibilidade de união ou de tipo entre R e S.

- **Ex:** construa uma operação para recuperar a lista dos nomes dos dependentes e os nomes de seus pais (tutores)

Tutor		
cpf	nome	data_nasc
000.111.222-33	Claudia	1980-10-05
222.333.444-55	Jose	1995-04-02
555.666.777-88	Amalia	1978-06-30
999.000.111-22	Cleber	1985-07-25

Dependente			
cpf_dep	nome_dep	data_nasc_dep	cpf_tutor
777.555.111-00	Armando	2000-02-25	000.111.222-33
666.444.222-44	Valeriano	2002-05-30	222.333.444-55
888.777.333-55	Cicrana	2001-07-21	222.333.444-55

Tutor		
cpf	nome	data_nasc
000.111.222-33	Claudia	1980-10-05
222.333.444-55	Jose	1995-04-02
555.666.777-88	Amalia	1978-06-30
999.000.111-22	Cleber	1985-07-25

Dependente			
cpf_dep	nome_dep	data_nasc_dep	cpf_tutor
777.555.111-00	Armando	2000-02-25	000.111.222-33
666.444.222-44	Valeriano	2002-05-30	222.333.444-55
888.777.333-55	Cicrana	2001-07-21	222.333.444-55

- Tutor X Dependente

cpf	nome	data_nasc	cpf_dep	nome_dep	data_nasc_dep	cpf_tutor
-----	------	-----------	---------	----------	---------------	-----------

Tutor		
cpf	nome	data_nasc
000.111.222-33	Claudia	1980-10-05
222.333.444-55	Jose	1995-04-02
555.666.777-88	Amalia	1978-06-30
999.000.111-22	Cleber	1985-07-25

Dependente			
cpf_dep	nome_dep	data_nasc_dep	cpf_tutor
777.555.111-00	Armando	2000-02-25	000.111.222-33
666.444.222-44	Valeriano	2002-05-30	222.333.444-55
888.777.333-55	Cicrana	2001-07-21	222.333.444-55

● Tutor X Dependente

cpf	nome	data_nasc	cpf_dep	nome_dep	data_nasc_dep	cpf_tutor
000.111.222-33	Claudia	1980-10-05	777.555.111-00	Armando	2000-02-25	000.111.222-33
222.333.444-55	Jose	1995-04-02	777.555.111-00	Armando	2000-02-25	000.111.222-33
555.666.777-88	Amalia	1978-06-30	777.555.111-00	Armando	2000-02-25	000.111.222-33
999.000.111-22	Cleber	1985-07-25	777.555.111-00	Armando	2000-02-25	000.111.222-33

Tutor		
cpf	nome	data_nasc
000.111.222-33	Claudia	1980-10-05
222.333.444-55	Jose	1995-04-02
555.666.777-88	Amalia	1978-06-30
999.000.111-22	Cleber	1985-07-25

Dependente			
cpf_dep	nome_dep	data_nasc_dep	cpf_tutor
777.555.111-00	Armando	2000-02-25	000.111.222-33
666.444.222-44	Valeriano	2002-05-30	222.333.444-55
888.777.333-55	Cicrana	2001-07-21	222.333.444-55

● Tutor X Dependente

cpf	nome	data_nasc	cpf_dep	nome_dep	data_nasc_dep	cpf_tutor
000.111.222-33	Claudia	1980-10-05	777.555.111-00	Armando	2000-02-25	000.111.222-33
222.333.444-55	Jose	1995-04-02	777.555.111-00	Armando	2000-02-25	000.111.222-33
555.666.777-88	Amalia	1978-06-30	777.555.111-00	Armando	2000-02-25	000.111.222-33
999.000.111-22	Cleber	1985-07-25	777.555.111-00	Armando	2000-02-25	000.111.222-33

Tutor		
cpf	nome	data_nasc
000.111.222-33	Claudia	1980-10-05
222.333.444-55	Jose	1995-04-02
555.666.777-88	Amalia	1978-06-30
999.000.111-22	Cleber	1985-07-25

Dependente			
cpf_dep	nome_dep	data_nasc_dep	cpf_tutor
777.555.111-00	Armando	2000-02-25	000.111.222-33
666.444.222-44	Valeriano	2002-05-30	222.333.444-55
888.777.333-55	Cicrana	2001-07-21	222.333.444-55

● Tutor X Dependente

cpf	nome	data_nasc	cpf_dep	nome_dep	data_nasc_dep	cpf_tutor
000.111.222-33	Claudia	1980-10-05	777.555.111-00	Armando	2000-02-25	000.111.222-33
222.333.444-55	Jose	1995-04-02	777.555.111-00	Armando	2000-02-25	000.111.222-33
555.666.777-88	Amalia	1978-06-30	777.555.111-00	Armando	2000-02-25	000.111.222-33
999.000.111-22	Cleber	1985-07-25	777.555.111-00	Armando	2000-02-25	000.111.222-33
000.111.222-33	Claudia	1980-10-05	666.444.222-44	Valeriano	2002-05-30	222.333.444-55
222.333.444-55	Jose	1995-04-02	666.444.222-44	Valeriano	2002-05-30	222.333.444-55
555.666.777-88	Amalia	1978-06-30	666.444.222-44	Valeriano	2002-05-30	222.333.444-55
999.000.111-22	Cleber	1985-07-25	666.444.222-44	Valeriano	2002-05-30	222.333.444-55

Tutor		
cpf	nome	data_nasc
000.111.222-33	Claudia	1980-10-05
222.333.444-55	Jose	1995-04-02
555.666.777-88	Amalia	1978-06-30
999.000.111-22	Cleber	1985-07-25

Dependente			
cpf_dep	nome_dep	data_nasc_dep	cpf_tutor
777.555.111-00	Armando	2000-02-25	000.111.222-33
666.444.222-44	Valeriano	2002-05-30	222.333.444-55
888.777.333-55	Cicrana	2001-07-21	222.333.444-55

● Tutor X Dependente

cpf	nome	data_nasc	cpf_dep	nome_dep	data_nasc_dep	cpf_tutor
000.111.222-33	Claudia	1980-10-05	777.555.111-00	Armando	2000-02-25	000.111.222-33
222.333.444-55	Jose	1995-04-02	777.555.111-00	Armando	2000-02-25	000.111.222-33
555.666.777-88	Amalia	1978-06-30	777.555.111-00	Armando	2000-02-25	000.111.222-33
999.000.111-22	Cleber	1985-07-25	777.555.111-00	Armando	2000-02-25	000.111.222-33
000.111.222-33	Claudia	1980-10-05	666.444.222-44	Valeriano	2002-05-30	222.333.444-55
222.333.444-55	Jose	1995-04-02	666.444.222-44	Valeriano	2002-05-30	222.333.444-55
555.666.777-88	Amalia	1978-06-30	666.444.222-44	Valeriano	2002-05-30	222.333.444-55
999.000.111-22	Cleber	1985-07-25	666.444.222-44	Valeriano	2002-05-30	222.333.444-55
000.111.222-33	Claudia	1980-10-05	888.777.333-55	Cicrana	2001-07-21	222.333.444-55
222.333.444-55	Jose	1995-04-02	888.777.333-55	Cicrana	2001-07-21	222.333.444-55
555.666.777-88	Amalia	1978-06-30	888.777.333-55	Cicrana	2001-07-21	222.333.444-55
999.000.111-22	Cleber	1985-07-25	888.777.333-55	Cicrana	2001-07-21	222.333.444-55

● Tutor X Dependente

cpf	nome	data_nasc	cpf_dep	nome_dep	data_nasc_dep	cpf_tutor
000.111.222-33	Claudia	1980-10-05	777.555.111-00	Armando	2000-02-25	000.111.222-33
222.333.444-55	Jose	1995-04-02	777.555.111-00	Armando	2000-02-25	000.111.222-33
555.666.777-88	Amalia	1978-06-30	777.555.111-00	Armando	2000-02-25	000.111.222-33
999.000.111-22	Cleber	1985-07-25	777.555.111-00	Armando	2000-02-25	000.111.222-33
000.111.222-33	Claudia	1980-10-05	666.444.222-44	Valeriano	2002-05-30	222.333.444-55
222.333.444-55	Jose	1995-04-02	666.444.222-44	Valeriano	2002-05-30	222.333.444-55
555.666.777-88	Amalia	1978-06-30	666.444.222-44	Valeriano	2002-05-30	222.333.444-55
999.000.111-22	Cleber	1985-07-25	666.444.222-44	Valeriano	2002-05-30	222.333.444-55
000.111.222-33	Claudia	1980-10-05	888.777.333-55	Cicrana	2001-07-21	222.333.444-55
222.333.444-55	Jose	1995-04-02	888.777.333-55	Cicrana	2001-07-21	222.333.444-55
555.666.777-88	Amalia	1978-06-30	888.777.333-55	Cicrana	2001-07-21	222.333.444-55
999.000.111-22	Cleber	1985-07-25	888.777.333-55	Cicrana	2001-07-21	222.333.444-55

**Observe que temos tuplas falsas agora.
Isto é, os tutores foram associados aos dependentes errados!**

“A tutora de *Armando* é o cpf *000.111.222-33* (*Claudia*)”

“O tutor de *Cicrana* é o cpf *222.333.444-55* (*Jose*)”

- Tutor X Dependente**

cpf	nome	data_nasc	cpf_dep	nome_dep	data_nasc_dep	cpf_tutor
000.111.222-33	Claudia	1980-10-05	777.555.111-00	Armando	2000-02-25	000.111.222-33
222.333.444-55	Jose	1995-04-02	777.555.111-00	Armando	2000-02-25	000.111.222-33
555.666.777-88	Amalia	1978-06-30	777.555.111-00	Armando	2000-02-25	000.111.222-33
999.000.111-22	Cleber	1985-07-25	777.555.111-00	Armando	2000-02-25	000.111.222-33
000.111.222-33	Claudia	1980-10-05	666.444.222-44	Valeriano	2002-05-30	222.333.444-55
222.333.444-55	Jose	1995-04-02	666.444.222-44	Valeriano	2002-05-30	222.333.444-55
555.666.777-88	Amalia	1978-06-30	666.444.222-44	Valeriano	2002-05-30	222.333.444-55
999.000.111-22	Cleber	1985-07-25	666.444.222-44	Valeriano	2002-05-30	222.333.444-55
000.111.222-33	Claudia	1980-10-05	888.777.333-55	Cicrana	2001-07-21	222.333.444-55
222.333.444-55	Jose	1995-04-02	888.777.333-55	Cicrana	2001-07-21	222.333.444-55
555.666.777-88	Amalia	1978-06-30	888.777.333-55	Cicrana	2001-07-21	222.333.444-55
999.000.111-22	Cleber	1985-07-25	888.777.333-55	Cicrana	2001-07-21	222.333.444-55

Para corrigir as tuplas falsas iremos usar o operador Seleção:

$$\sigma_{\text{cpf}=\text{cpf_tutor}}(\text{Tutor X Dependente})$$

- $\sigma_{\text{cpf}=\text{cpf_tutor}}$ (Tutor X Dependente)

cpf	nome	data_nasc	cpf_dep	nome_dep	data_nasc_dep	cpf_tutor
000.111.222-33	Claudia	1980-10-05	777.555.111-00	Armando	2000-02-25	000.111.222-33
222.333.444-55	Jose	1995-04-02	666.444.222-44	Valeriano	2002-05-30	222.333.444-55
222.333.444-55	Jose	1995-04-02	888.777.333-55	Cicrana	2001-07-21	222.333.444-55

Agora temos os tutores associados aos dependentes corretos.

Operação Produto Cartesiano (**X**)

- De maneira formal:
 - O resultado de $\mathbf{R} (A_1, A_2, \dots, A_n) \mathbf{X} \mathbf{S} (B_1, B_2, \dots, B_m)$ é uma relação \mathbf{Q} com $n+m$ atributos $\mathbf{Q} (A_1, A_2, \dots, A_n, B_1, B_2, \dots, B_m)$ nessa ordem
- A relação \mathbf{Q} possui uma tupla para cada combinação de tuplas \mathbf{R} e \mathbf{S}
 - Se \mathbf{R} tem \mathbf{t} tuplas e \mathbf{S} tem \mathbf{z} tuplas então $\mathbf{Q} = \mathbf{R} \mathbf{x} \mathbf{S}$ possui $\mathbf{t} * \mathbf{z}$ tuplas

$$\mathbf{Q} = \mathbf{R} \mathbf{X} \mathbf{S}$$

Atividade - Produto cartesiano

- Sejam as tabelas abaixo, contendo os seguintes atributos:
 - **Produtos** (id_prod: **int+PK**, nome_prod: **varchar(20)**, valor: **double**, qtd: **int**, id_fornec: **int+FK**)
 - **Fornecedor** (id_forn: **int+PK**, nome_forn: **varchar(20)**, area_atuacao: **varchar(30)**, cnae: **int**)
- Realize uma consulta usando as operações seleção, projeção, renomear e/ou produto cartesiano para:
 - Encontrar o nome dos fornecedores que fabricam o produto de nome **"tinta branco fosco"**

Problema da Operação Produto Cartesiano

- A operação produto cartesiano gera tabelas muito grandes
 - Desempenho da operação produto cartesiano é ruim
- A operação produto cartesiano requer o uso de operações Seleção para remover tuplas falsas
- **Será que conseguimos melhorar esses problemas?**

Operação Junção (\bowtie)

- Operação que permite combinar tuplas (de maneira semelhante ao produto cartesiano), utilizando para isso uma condição

- **Entrada:**

- relação $R(A_1, A_2, \dots, A_n)$
- relação $S(B_1, B_2, \dots, B_m)$
- condição de seleção

$$Q = R \bowtie_{\text{condicao_de_selecao}}(S)$$

- **Saída:**

- relação $Q(A_1, A_2, \dots, A_n, B_1, B_2, \dots, B_m)$

Operação Junção (\bowtie)

- Essa operação tem como saída (**Q**) tuplas de **R** e **S** combinadas que satisfazem a condição de seleção
 - Tuplas cujos atributos são **NULL** ou condição de seleção seja **FALSE** não aparecem no resultado **Q**

Tutor		
cpf	nome	data_nasc
000.111.222-33	Claudia	1980-10-05
222.333.444-55	Jose	1995-04-02
555.666.777-88	Amalia	1978-06-30
999.000.111-22	Cleber	1985-07-25

Dependente			
cpf_dep	nome_dep	data_nasc_dep	cpf_tutor
777.555.111-00	Armando	2000-02-25	000.111.222-33
666.444.222-44	Valeriano	2002-05-30	222.333.444-55
888.777.333-55	Cicrana	2001-07-21	222.333.444-55

- Tutor $\bowtie_{\text{cpf} = \text{cpf_tutor}}$ (Dependente)

Tutor		
cpf	nome	data_nasc
000.111.222-33	Claudia	1980-10-05
222.333.444-55	Jose	1995-04-02
555.666.777-88	Amalia	1978-06-30
999.000.111-22	Cleber	1985-07-25

Dependente			
cpf_dep	nome_dep	data_nasc_dep	cpf_tutor
777.555.111-00	Armando	2000-02-25	000.111.222-33
666.444.222-44	Valeriano	2002-05-30	222.333.444-55
888.777.333-55	Cicrana	2001-07-21	222.333.444-55

- Tutor \bowtie (Dependente)
 $\text{cpf} = \text{cpf_tutor}$

Tutor		
cpf	nome	data_nasc
000.111.222-33	Claudia	1980-10-05
222.333.444-55	Jose	1995-04-02
555.666.777-88	Amalia	1978-06-30
999.000.111-22	Cleber	1985-07-25

Dependente			
cpf_dep	nome_dep	data_nasc_dep	cpf_tutor
777.555.111-00	Armando	2000-02-25	000.111.222-33
666.444.222-44	Valeriano	2002-05-30	222.333.444-55
888.777.333-55	Cicrana	2001-07-21	222.333.444-55

- Tutor \bowtie _{cpf = cpf_tutor} (Dependente)

cpf	nome	data_nasc	cpf_dep	nome_dep	data_nasc_dep	cpf_tutor
000.111.222-33	Claudia	1980-10-05	777.555.111-00	Armando	2000-02-25	000.111.222-33
222.333.444-55	Jose	1995-04-02	666.444.222-44	Valeriano	2002-05-30	222.333.444-55
222.333.444-55	Jose	1995-04-02	888.777.333-55	Cicrana	2001-07-21	222.333.444-55

Tutor		
cpf	nome	data_nasc
000.111.222-33	Claudia	1980-10-05
222.333.444-55	Jose	1995-04-02
555.666.777-88	Amalia	1978-06-30
999.000.111-22	Cleber	1985-07-25

Dependente			
cpf_dep	nome_dep	data_nasc_dep	cpf_tutor
777.555.111-00	Armando	2000-02-25	000.111.222-33
666.444.222-44	Valeriano	2002-05-30	222.333.444-55
888.777.333-55	Cicrana	2001-07-21	222.333.444-55

- $\text{Tutor} \bowtie_{\text{cpf} = \text{cpf_tutor}} (\text{Dependente}) = \sigma_{\text{cpf}=\text{cpf_tutor}} (\text{Tutor X Dependente})$

cpf	nome	data_nasc	cpf_dep	nome_dep	data_nasc_dep	cpf_tutor
000.111.222-33	Claudia	1980-10-05	777.555.111-00	Armando	2000-02-25	000.111.222-33
222.333.444-55	Jose	1995-04-02	666.444.222-44	Valeriano	2002-05-30	222.333.444-55
222.333.444-55	Jose	1995-04-02	888.777.333-55	Cicrana	2001-07-21	222.333.444-55

Operação Divisão (\div)

$$Q = R \div S$$

- Operação binária cujo resultado contém apenas os elementos de **R** que se relacionam com TODOS os elementos de **S**

Estoque	
Produto	Preco
Ameixa	5,2
Pacoquita	2,1
Uva	10,5
Pasta de Dente	3,5

Preco_Produtos	
Preco	
	5,2
	2,1
	3,5

- Estoque ÷ Preco_Produtos

Produto
Ameixa

Estoque	
Produto	Preco
Ameixa	5,2
Pacoquita	2,1
Uva	10,5
Pasta de Dente	3,5

Preco_Produtos	
Preco	
	5,2
	2,1
	3,5

- Estoque ÷ Preco_Produtos

Produto
Ameixa
Pacoquita

Estoque	
Produto	Preco
Ameixa	5,2
Pacoquita	2,1
Uva	10,5
Pasta de Dente	3,5

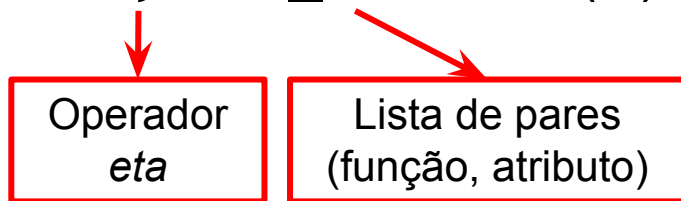
Preco_Produtos
Preco
5,2
2,1
3,5

- Estoque ÷ Preco_Produtos

Produto
Ameixa
Pacoquita
Pasta de Dente

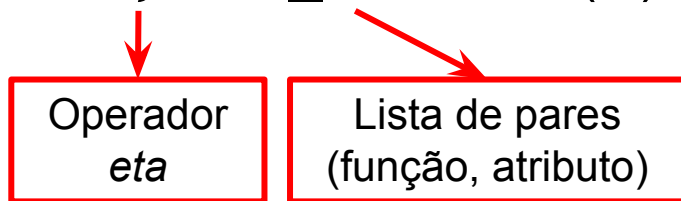
Funções de agregação

- Funções de agregação fazem um “resumo” de um conjunto de valores de tuplas do banco de dados
 - SOMA, MÉDIA, MÁXIMO, MÍNIMO, CONTA
 - SUM, AVG, MAX, MIN, COUNT
- **Formato:** $Q = \xi \langle \text{lista_funcoes} \rangle (R)$



Funções de agregação

- Funções de agregação fazem um “resumo” de um conjunto de valores de tuplas do banco de dados
 - SOMA, MÉDIA, MÁXIMO, MÍNIMO, CONTA
 - SUM, AVG, MAX, MIN, COUNT
- **Formato:** $Q = \xi \langle \text{lista_funcoes} \rangle (R)$



Os atributos de **Q** terão o seguinte nome: nomeFuncao_atributoUtilizado

FUNCIONARIO

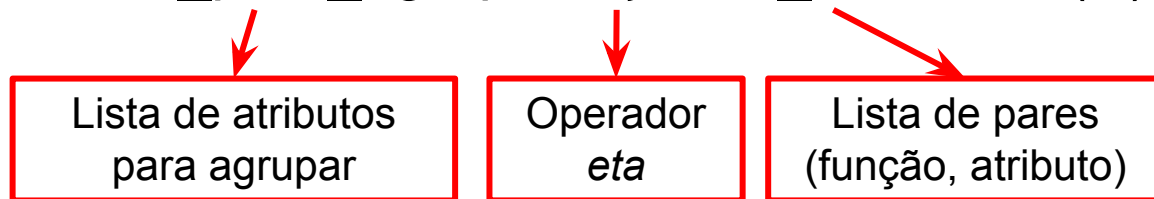
Pnome	Minicial	Unome	Cpf	Datanasc	Endereco	Sexo	Salario
João	B	Silva	12345678966	09-01-1965	Rua das Flores, 751, São Paulo, SP	M	30.000
Fernando	T	Wong	33344555587	08-12-1955	Rua da Lapa, 34, São Paulo, SP	M	40.000
Alice	J	Zelaya	99988777767	19-01-1968	Rua Souza Lima, 35, Curitiba, PR	F	25.000
Jennifer	S	Souza	98765432168	20-06-1941	Av. Arthur de Lima, 54, Santo André, SP	F	43.000
Ronaldo	K	Lima	66688444476	15-09-1962	Rua Rebouças, 65, Piracicaba, SP	M	38.000
Joice	A	Leite	45345345376	31-07-1972	Av. Lucas Obes, 74, São Paulo, SP	F	25.000
André	V	Pereira	98798798733	29-03-1969	Rua Timbira, 35, São Paulo, SP	M	25.000
Jorge	E	Brito	88866555576	10-11-1937	Rua do Horto, 35, São Paulo, SP	M	55.000

- $\xi < \text{CONTE}(\text{Cpf}), \text{MEDIA}(\text{Salario}) > (\text{FUNCIONARIO})$

Contador_cpf	Media_salario
8	35.125

Funções de agrupamento

- Funções de agrupamento juntam um conjunto de valores de por categoria, e depois executam a função de agregação
- **Formato:** $Q = \langle \text{atributos_para_agrupar} \rangle \xi \langle \text{lista_funcoes} \rangle (R)$



FUNCIONARIO

Pnome	Minicial	Unome	Cpf	Datanasc	Endereco	Sexo	Salario	Cpf_supervisor	Dnr
João	B	Silva	12345678966	09-01-1965	Rua das Flores, 751, São Paulo, SP	M	30.000	33344555587	5
Fernando	T	Wong	33344555587	08-12-1955	Rua da Lapa, 34, São Paulo, SP	M	40.000	88866555576	5
Alice	J	Zelaya	99988777767	19-01-1968	Rua Souza Lima, 35, Curitiba, PR	F	25.000	98765432168	4
Jennifer	S	Souza	98765432168	20-06-1941	Av. Arthur de Lima, 54, Santo André, SP	F	43.000	88866555576	4
Ronaldo	K	Lima	66688444476	15-09-1962	Rua Rebouças, 65, Piracicaba, SP	M	38.000	33344555587	5
Joice	A	Leite	45345345376	31-07-1972	Av. Lucas Obes, 74, São Paulo, SP	F	25.000	33344555587	5
André	V	Pereira	98798798733	29-03-1969	Rua Timbira, 35, São Paulo, SP	M	25.000	98765432168	4
Jorge	E	Brito	88866555576	10-11-1937	Rua do Horto, 35, São Paulo, SP	M	55.000	NULL	1

- <Dnr> ξ <**CONTE**(Cpf), **MEDIA**(Salario)> (FUNCIONARIO)

Dnr	Contador_cpf	Media_salario
5	4	33.250
4	3	31.000
1	1	55.000

FUNCIONARIO

Pnome	Minicial	Unome	Cpf	Datanasc	Endereco	Sexo	Salario	Cpf_supervisor	Dnr
João	B	Silva	12345678966	09-01-1965	Rua das Flores, 751, São Paulo, SP	M	30.000	33344555587	5
Fernando	T	Wong	33344555587	08-12-1955	Rua da Lapa, 34, São Paulo, SP	M	40.000	88866555576	5
Alice	J	Zelaya	99988777767	19-01-1968	Rua Souza Lima, 35, Curitiba, PR	F	25.000	98765432168	4
Jennifer	S	Souza	98765432168	20-06-1941	Av. Arthur de Lima, 54, Santo André, SP	F	43.000	88866555576	4
Ronaldo	K	Lima	66688444476	15-09-1962	Rua Rebouças, 65, Piracicaba, SP	M	38.000	33344555587	5
Joice	A	Leite	45345345376	31-07-1972	Av. Lucas Obes, 74, São Paulo, SP	F	25.000	33344555587	5
André	V	Pereira	98798798733	29-03-1969	Rua Timbira, 35, São Paulo, SP	M	25.000	98765432168	4
Jorge	E	Brito	88866555576	10-11-1937	Rua do Horto, 35, São Paulo, SP	M	55.000	NULL	1

- $\langle \text{Dnr} \rangle \xi \langle \text{CONTE}(\text{Cpf}), \text{MEDIA}(\text{Salario}) \rangle (\text{FUNCIONARIO})$

Dnr	Contador_cpf	Media_salario
5	4	33.250
4	3	31.000
1	1	55.000

No depto número 5, existem 4 CPFs, e a média de salário desses funcionários é de 33,25

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.