



# Disciplina: Fundamentos de Banco de Dados

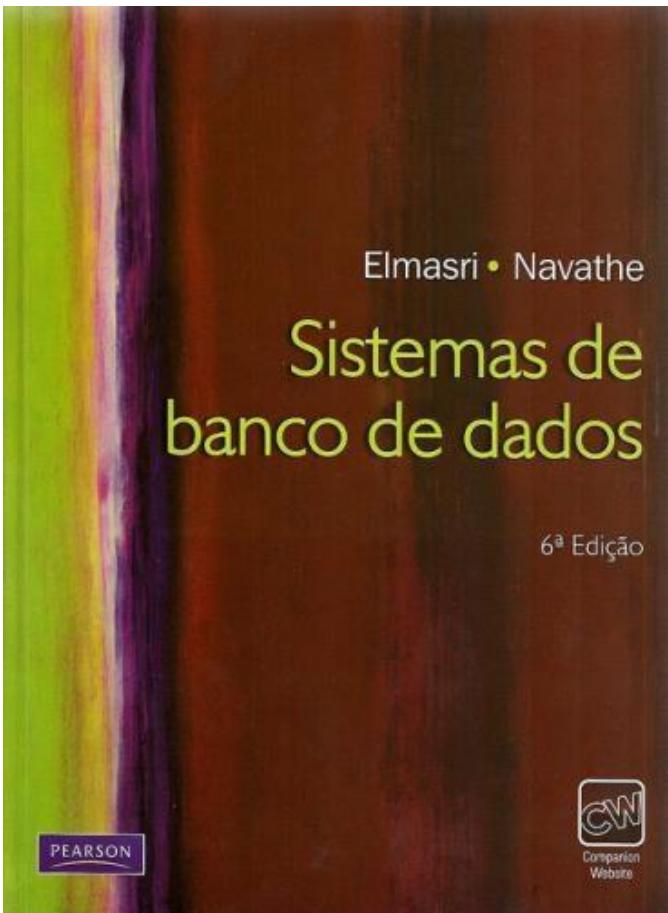
## Parte 4 SQL - Funções de Agregação e Agrupamento

---

Professora: Marília S. Mendes  
E-mail: marilia.mendes@ufc.br

# Conteúdo desta aula:

---



- Capítulo 8: SQL-99
  - Seção 8.5.7: Funções Agregadas em SQL
  - Seção 8.5.8: Agrupamento - as cláusulas GROUP BY e HAVING

# Organização da aula

---

→ **Funções de Agregação**

→ Agrupamento



# Agregação

---

- ➔ Funções que atuam em **coleções** de valores do banco de dados
- ➔ Usadas em **consultas estatísticas** simples que resumem informações do banco de dados
- ➔ Exemplos:
  - Calcular a média dos salários dos empregados
  - Calcular quantos empregados a empresa possui



# Funções de Agregação

---

→ COUNT

→ SUM

→ MAX

→ MIN

→ AVG



# COUNT

---

Função que **conta** o número de  
tuplas retornadas por uma consulta



# COUNT - Exemplo

```
SELECT *  
FROM EMPREGADO
```

	PNAME	MINICIAL	UNOME	SSN	DATANASC	ENDERECO	SEXO	SALARIO	SUPERSSN	DNO
▶	John	B	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	M	30000	333445555	5
	Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	40000	888665555	5
	Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5
	Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	M	38000	333445555	5
	James	E	Borg	888665555	1937-11-10	450 Stone, Houston, TX	M	55000		1
	Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4
	Ahmad	V	Jabbar	987987987	1969-03-29	980 Dallas, Houston, TX	M	25000	987654321	4
	Alicia	J	Zelaya	999887777	1968-01-19	3321 Castle, Spring, TX	F	25000	987654321	4

# COUNT - Exemplo

```
SELECT COUNT(*)  
FROM EMPREGADO
```

	COUNT(*)
▶	8

```
SELECT *  
FROM EMPREGADO
```

PNAME	MINITIAL	UNOME	SSN	DATANASC	ENDERECO	SEXO	SALARIO	SUPERSSN	DNO
John	B	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	40000	888665555	5
Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5
Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	M	38000	333445555	5
James	E	Borg	888665555	1937-11-10	450 Stone, Houston, TX	M	55000		1
Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4
Ahmad	V	Jabbar	987987987	1969-03-29	980 Dallas, Houston, TX	M	25000	987654321	4
Alicia	J	Zelaya	999887777	1968-01-19	3321 Castle, Spring, TX	F	25000	987654321	4

8 tuplas

# COUNT - Exemplo

---

```
SELECT COUNT(SSN)  
FROM EMPREGADO
```

	COUNT(SSN)
▶	8



# COUNT - Exemplo

```
SELECT COUNT(SSN)  
FROM EMPREGADO
```

COUNT(SSN)
8

```
SELECT SSN  
FROM EMPREGADO
```

SSN
123456789
333445555
453453453
666884444
888665555
987654321
987987987
999887777

8 tuplas



# Nomeando a coluna

---

```
SELECT COUNT(SSN)  
FROM EMPREGADO
```

COUNT(SSN)
► 8

```
SELECT COUNT(SSN) AS NUMEMPREGADOS  
FROM EMPREGADO
```

NUMEMPREGADOS
► 8



# Contando ocorrências de valores

---

## → COUNT

- Contagem de **tuplas**
- E se quisermos contar quantos salários diferentes existem no banco de dados?
- Contagem de valores



# Contando ocorrências de valores

---

## → Solução

- Transformar **contagem de valores** em **contagem de tuplas**
- Elaborar consulta que seleciona todos os valores distintos de salário, e aplicar COUNT



# Exemplo

---

```
SELECT DISTINCT SALARIO  
FROM EMPREGADO
```

SALARIO
30000
40000
25000
38000
55000
43000



# Exemplo

---

```
SELECT COUNT(DISTINCT SALARIO)  
      AS NUMSALARIOS  
  FROM EMPREGADO
```

	NUMSALARIOS
▶	6



## COUNT(DISTINCT Salario) x COUNT(Salario)

- Se não usarmos DISTINCT, obtemos o mesmo resultado de **COUNT(\*)**
  - Resumindo: **COUNT(SALARIO) = COUNT(\*)**
  - Exceção: valores *null* em Salario
- O DISTINCT garante que só serão contados os valores diferentes de salário

# Consultas podem ser mais complexas

---

→ Contar número de empregados do departamento "Pesquisa"

	NUMEMP
►	4



# SUM

---

Função aplicada a um conjunto  
de valores que retorna a **soma**  
de tais valores



# SUM - Exemplo

Calcular a soma dos salários dos empregados

PNAME	MINICIAL	UNOME	SSN	DATANASC	ENDERECO	SEXO	SALARIO	SUPERSSN	DNO
John	B	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	40000	888665555	5
Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5
Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	M	38000	333445555	5
James	E	Borg	888665555	1937-11-10	450 Stone, Houston, TX	M	55000		1
Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4
Ahmad	V	Jabbar	987987987	1969-03-29	980 Dallas, Houston, TX	M	25000	987654321	4
Alicia	J	Zelaya	999887777	1968-01-19	3321 Castle, Spring, TX	F	25000	987654321	4

281000

## SUM - Exemplo

---

```
SELECT SUM(SALARIO)  
FROM EMPREGADO
```

	SUM(SALARIO)
▶	281000



# SUM - Exemplo

---



Calcular a soma dos salários dos empregados que trabalham no projeto "ProdutoX"

SOMASALARIO
55000



# MAX e MIN

---

**MAX:** Função aplicada a um conjunto  
de valores que retorna o **valor**  
**máximo** do conjunto

**MIN:** Função aplicada a um conjunto  
de valores que retorna o **valor**  
**mínimo** do conjunto



# MAX - Exemplo

---

→ Retornar o maior salário de empregado

```
SELECT MAX(SALARIO)  
FROM EMPREGADO
```

MAX(SALARIO)
55000



# MIN - Exemplo

---

→ Retornar o número mínimo de horas em que um empregado trabalha em um projeto

	MIN(HORAS)
►	5



# AVG

---

Função aplicada a um conjunto  
de valores que retorna a  
**média** dos valores do conjunto



# AVG - Exemplo



Calcular a média dos salários dos empregados

PNAME	MINICIAL	UNOME	SSN	DATANASC	ENDERECO	SEXO	SALARIO	SUPERSSN	DNO
John	B	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	40000	888665555	5
Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5
Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	M	38000	333445555	5
James	E	Borg	888665555	1937-11-10	450 Stone, Houston, TX	M	55000		1
Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4
Ahmad	V	Jabbar	987987987	1969-03-29	980 Dallas, Houston, TX	M	25000	987654321	4
Alicia	J	Zelaya	999887777	1968-01-19	3321 Castle, Spring, TX	F	25000	987654321	4

$$281000 \div 8 = 35125$$

Soma dos salários

Número de empregados

## AVG - Exemplo

---

```
SELECT AVG(SALARIO)  
FROM EMPREGADO
```

	AVG(SALARIO)
▶	35125



# Misturando funções

---

```
SELECT SUM(SALARIO), MAX(SALARIO), MIN(SALARIO), AVG(SALARIO)  
FROM EMPREGADO
```

	SUM(SALARIO)	MAX(SALARIO)	MIN(SALARIO)	AVG(SALARIO)
▶	281000	55000	25000	35125



## Até agora...

---

- Exemplos simples de como obter informações resumidas do banco de dados
- Na maioria das vezes precisaremos usar estas informações como critérios de consultas para obter resultados mais elaborados
  - Consultas aninhadas



# Consultas Aninhadas usando Agregação

- Recuperar os nomes de todos os empregados que possuem mais de 2 dependentes

```
SELECT PNOME, UNOME  
FROM EMPREGADO E  
WHERE (SELECT COUNT(*)  
      FROM DEPENDENTE D  
      WHERE E.SSN=D.ESSN) >= 2
```

	PNOME	UNOME
▶	John	Smith
	Franklin	Wong

# Organização da aula

---

- Funções de Agregação
- Agrupamento



# Sub-Grupos

---

→ Exemplos de consultas desejáveis:

- Encontrar a média dos salários dos empregados em cada departamento
- Contar os empregados que trabalham em cada projeto

→ Isso exige: aplicar funções de agregação a **sub-grupos de tuplas** de uma tabela, **baseado no valor de determinado atributo**



# Sub-Grupos

---

## → Procedimento:

- Agrupar tuplas que possuem o mesmo valor de algum atributo
- Aplicar a função de agregação a cada grupo de tuplas



# Exemplo



Encontrar a média dos salários dos empregados em cada departamento



PNAME	MINITIAL	UNOME	SSN	DATANASC	ENDERECO	SEXO	SALARIO	SUPERSSN	DNO
John	B	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	40000	888665555	5
Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5
Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	M	38000	333445555	5
James	E	Borg	888665555	1937-11-10	450 Stone, Houston, TX	M	55000		1
Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4
Ahmad	V	Jabbar	987987987	1969-03-29	980 Dallas, Houston, TX	M	25000	987654321	4
Alicia	J	Zelaya	999887777	1968-01-19	3321 Castle, Spring, TX	F	25000	987654321	4

# Exemplo



Encontrar a média dos salários dos empregados em cada departamento

PNAME	MINITIAL	UNOME	SSN	DATANASC	ENDERECO	SEXO	SALARIO	SUPERSSN	DNO
John	B	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	40000	888665555	5
Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5
Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	M	38000	333445555	5
James	E	Borg	888665555	1937-11-10	450 Stone, Houston, TX	M	55000		1
Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4
Ahmad	V	Jabbar	987987987	1969-03-29	980 Dallas, Houston, TX	M	25000	987654321	4
Alicia	J	Zelaya	999887777	1968-01-19	3321 Castle, Spring, TX	F	25000	987654321	4

# Exemplo

Encontrar a média dos salários dos empregados em cada departamento

PNAME	MINITIAL	UNOME	SSN	DATANASC	ENDERECO	SEXO	SALARIO	SUPERSSN	DNO
John	B	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	40000	888665555	5
Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5
Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	M	38000	333445555	5
James	E	Borg	888665555	1937-11-10	450 Stone, Houston, TX	M	55000		1
Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4
Ahmad	V	Jabbar	987987987	1969-03-29	980 Dallas, Houston, TX	M	25000	987654321	4
Alicia	J	Zelaya	999887777	1968-01-19	3321 Castle, Spring, TX	F	25000	987654321	4

$$133000 \div 4 = 33250$$

Soma dos salários

Número de empregados

# Exemplo



Encontrar a média dos salários dos empregados em cada departamento

PNAME	MINITIAL	UNOME	SSN	DATANASC	ENDERECO	SEXO	SALARIO	SUPERSSN	DNO
John	B	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	40000	888665555	5
Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5
Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	M	38000	333445555	5
James	E	Borg	888665555	1937-11-10	450 Stone, Houston, TX	M	55000		1
Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4
Ahmad	V	Jabbar	987987987	1969-03-29	980 Dallas, Houston, TX	M	25000	987654321	4
Alicia	J	Zelaya	999887777	1968-01-19	3321 Castle, Spring, TX	F	25000	987654321	4

$$55000 \div 1 = 55000$$

Soma dos salários

Número de empregados

# Exemplo



Encontrar a média dos salários dos empregados em cada departamento

PNAME	MNICIAL	UNOME	SSN	DATANASC	ENDERECO	SEXO	SALARIO	SUPERSSN	DNO
John	B	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	40000	888665555	5
Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5
Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	M	38000	333445555	5
James	E	Borg	888665555	1937-11-10	450 Stone, Houston, TX	M	55000		1
Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4
Ahmad	V	Jabbar	987987987	1969-03-29	980 Dallas, Houston, TX	M	25000	987654321	4
Alicia	J	Zelaya	999887777	1968-01-19	3321 Castle, Spring, TX	F	25000	987654321	4

$$93000 \div 3 = 31000$$

Soma dos salários

Número de empregados

# Exemplo

→ Encontrar a média dos salários dos empregados em cada departamento

```
SELECT DNO, AVG(SALARIO)  
FROM EMPREGADO E  
GROUP BY DNO
```

	DNO	AVG(SALARIO)
▶	1	55000
	4	31000
	5	33250

# GROUP BY

→ As colunas agrupadas **precisam aparecer** no SELECT

```
SELECT DNO, AVG(SALARIO)  
FROM EMPREGADO E  
GROUP BY DNO
```

CORRETO

```
SELECT AVG(SALARIO)  
FROM EMPREGADO E  
GROUP BY DNO
```

INCORRETO

## GROUP BY - Exemplo

---

- Para cada projeto, recuperar o número do projeto, o nome do projeto e o número de empregados que trabalham no projeto



## GROUP BY - Exemplo

```
SELECT PNUMERO, PJNOME, COUNT(*) AS NUMEMP  
FROM PROJETO, TRABALHA_EM  
WHERE PNUMERO=PNO  
GROUP BY PNUMERO, PJNOME
```

	PNUMERO	PJNOME	NUMEMP
▶	1	ProdutoX	2
	2	ProdutoY	3
	3	ProdutoZ	2
	10	Automatização	3
	20	Reorganização	3
	30	Novos Benefícios	3

# GROUP BY - Exemplo

---

→ Exemplo mostra

- Possível usar junções em conjunto com GROUP BY
- Ordem de aplicação das operações:
  1. Junção
  2. Agrupamento
  3. Agregação



# HAVING

---

- Às vezes é necessário aplicar operações de agregação **apenas a grupos que satisfazem a certas condições**
- Exemplo: Retornar o nome, número do projeto e o número de empregados que trabalham no projeto, mas apenas para os projetos no qual trabalham mais de 2 empregados
- Solução: cláusula **HAVING**
  - usada junto com GROUP BY



## HAVING – Exemplo

```
SELECT PNUMERO, PJNOME, COUNT(*) AS NUMEMP  
FROM PROJETO, TRABALHA_EM  
WHERE PNUMERO=PNO  
GROUP BY PNUMERO, PJNOME  
HAVING COUNT(*) > 2
```

	PNUMERO	PJNOME	NUMEMP
▶	2	Produto Y	3
	10	Automatização	3
	20	Reorganização	3
	30	Novos Benefícios	3



## HAVING x WHERE

---

- WHERE: limita as tuplas sobre as quais as funções de agregação serão aplicadas
- HAVING: serve para selecionar grupos inteiros



## Exemplo incorreto

→ Contar o número total de empregados cujos salários excedem 40000 em cada departamento, mas somente para os departamentos onde trabalham mais de 5 empregados

```
SELECT DNOME, COUNT(*)  
FROM DEPARTAMENTO, EMPREGADO  
WHERE DNUMERO=DNO AND SALARIO > 40000  
GROUP BY DNOME  
HAVING COUNT(*) > 5
```

## Exemplo incorreto

---

- Esta consulta seleciona os departamentos que possuem mais de 5 empregados que ganham mais do que 40000
- O WHERE é aplicado primeiro, e seleciona somente empregados que ganham mais de 40000

```
SELECT DNOME, COUNT(*)  
FROM DEPARTAMENTO, EMPREGADO  
WHERE DNUMERO=DNO AND SALARIO > 40000  
GROUP BY DNOME  
HAVING COUNT(*) > 5
```



## Exemplo CORRETO

---

→ Usar consulta aninhada

```
SELECT DNUMERO, COUNT(*)  
FROM DEPARTAMENTO, EMPREGADO  
WHERE DNUMERO=DNO AND SALARIO > 40000  
AND DNUMERO IN  
  (SELECT DNO FROM EMPREGADO  
   GROUP BY DNO  
   HAVING COUNT(*) > 5)  
GROUP BY DNUMERO
```



## Exercício 1

---

→ Selecionar o nome e número dos projetos cuja soma de horas de empregados alocados é superior a 50



# Exercício 1 - Resposta

---

- Selecionar o nome e número dos projetos cuja soma de horas de empregados alocados é superior a 50

```
SELECT PNO, PJNOME  
FROM PROJETO, TRABALHA_EM  
WHERE PNO=PNUMERO  
GROUP BY PNO, PJNOME  
HAVING SUM(Horas) > 50
```



## Exercício 2

---

- Selecionar os nomes dos departamentos que possuem mais de uma localidade



## Exercício 2 - Resposta

---

→ Selecionar os nomes dos departamentos que possuem mais de uma localidade

```
SELECT DNOME  
FROM DEPARTAMENTO D, DEPT_LOCALIZACOES I  
WHERE D.DNUMERO=I.DNUMERO  
GROUP BY DNOME  
HAVING COUNT(*) > 1
```



## Exercício 3

---

→ Para cada projeto, selecionar a média de horas em que empregados estão alocados



## Exercício 3 - Resposta

---

→ Para cada projeto, selecionar a média de horas em que empregados estão alocados

```
SELECT PNO, AVG(HORAS)  
FROM TRABALHA_EM  
GROUP BY PNO
```



# Bibliografia Utilizada nesta aula

---

- ▶ ELMASRI, R.; NAVATHE, S. B. Sistemas de banco de dados. 6 ed. Pearson/Addison-Wesley, 2011. ISBN: 9788579360855
  
- ▶ OLIVEIRA, C.H. SQL: Curso prático. Novatec, 2002. ISBN: 9788575220245 .