



SQL-Agregando Dados usando funções de grupo

Aula 4 - Simone Dominico



Funções de Agregação

Operam em conjuntos de linhas para fornecer um resultado por grupo, ignoram valores nulos com exceção do COUNT

Funções de Agregação

- Maior tamanho na tabela part

```
tpch=> SELECT MAX(p_size)  
FROM part;
```


Funções de Agregação

- Maior tamanho na tabela part

```
tpch=> SELECT MAX(p_size)
FROM part;
```

```
max
-----
50
(1 row)
```


Funções de Agregação

- *AVG*
- *COUNT*
- *MAX*
- *MIN*
- *SUM*
- *VARIANCE*

Funções de Agregação

```
SELECT coluna,group_function(coluna)  
FROM tabela  
WHERE condicao  
GROUP BY coluna1  
ORDER BY coluna2
```


Funções de Agregação - AVG

```
tpch=> SELECT AVG(p_size)  
FROM part;
```

SOMENTE DADOS NUMÉRICOS

Funções de Agregação - AVG

```
tpch=> SELECT AVG(p_size)  
FROM part;
```

SOMENTE DADOS NUMÉRICOS

```
      avg  
-----  
25.4271050000000000  
(1 row)
```


Funções de Agregação - SUM

```
tpch=> SELECT SUM(p_size)  
FROM part;
```

SOMENTE DADOS NUMÉRICOS

```
sum  
-----  
5085421  
(1 row)
```


Funções de Agregação - VARIANCE

```
tpch=> SELECT VARIANCE(c.c_acctbal)  
FROM customer c, nation n  
WHERE c.c_acctbal>2000;
```

SOMENTE DADOS NUMÉRICOS

Funções de Agregação - VARIANCE

```
tpch=> SELECT VARIANCE(c.c_acctbal)
FROM customer c, nation n
WHERE c.c_acctbal>2000;
```

SOMENTE DADOS NUMÉRICOS

```
      variance
-----
5353870.127608505224
(1 row)
```


Funções de Agregação - MIN E MAX

```
tpch=> SELECT MIN(p_size)  
FROM part;
```

QUALQUER TIPO DE DADOS

Funções de Agregação - MIN E MAX

```
tpch=> SELECT MIN(p_size)
FROM part;
```

QUALQUER TIPO DE DADOS

```
min
-----
      1
(1 row)
```


Funções de Agregação - MIN E MAX

```
tpch=> SELECT MIN(l_shipdate), MAX(l_shipdate)
FROM lineitem;
```

QUALQUER TIPO DE DADOS

min		max
-----+-----		
1992-01-02		1998-12-01
(1 row)		

Funções de Agregação - MIN E MAX

```
tpch=> SELECT MIN(r_name), MAX(r_name)
FROM region;
```

QUALQUER TIPO DE DADOS

min

|

max

+

AFRICA

|

MIDDLE EAST

(1 row)

Funções de Agregação - COUNT

```
tpch=> SELECT count(*)  
FROM lineitem;
```

Número de linhas de uma tabela

```
count  
-----  
6001215  
(1 row)
```


Funções de Agregação - COUNT

```
tpch=> SELECT COUNT(DISTINCT(n_regionkey))  
FROM nation;
```

Regiões distintas na tabela Nation

```
count  
-----  
      5  
(1 row)
```


Criando Grupo de Dados - GROUP BY

Ao utilizar a cláusula GROUP BY dividimos os registros que serão agregados em grupos de valores.

Criando Grupo de Dados - GROUP BY

```
tpch=> select c_mktsegment from customer  
tpch-> GROUP BY c_mktsegment;
```


Criando Grupo de Dados - GROUP BY

```
tpch=> select c_mktsegment from customer  
tpch-> GROUP BY c_mktsegment;
```

```
c_mktsegment  
-----  
FURNITURE  
MACHINERY  
AUTOMOBILE  
BUILDING  
HOUSEHOLD  
(5 rows)
```


Criando Grupo de Dados - GROUP BY

```
tpch=> select c_name, c_mktsegment from customer  
GROUP BY c_mktsegment;
```


Criando Grupo de Dados - GROUP BY

```
tpch=> select c_name, c_mktsegment from customer  
GROUP BY c_mktsegment;
```

```
ERRO:  coluna "customer.c_name" deve aparecer na cláusula GROUP BY ou se  
r utilizada em uma função de agregação
```

```
LINE 1: select c_name, c_mktsegment from customer
```


Agrupando por mais de Uma Coluna

```
tpch=> SELECT n.n_name, c.c_mktsegment  
FROM customer c, nation n  
GROUP BY c.c_mktsegment, n.n_name;
```


Agrupando por mais de Uma Coluna

```
tpch=> SELECT n.n_name, c.c_mktsegment
FROM customer c, nation n
GROUP BY c.c_mktsegment, n.n_name;
```

n_name	c_mktsegment
MOZAMBIQUE	AUTOMOBILE
ETHIOPIA	FURNITURE
MOROCCO	AUTOMOBILE
CHINA	HOUSEHOLD
BRAZIL	BUILDING

Consultas Ilegais - Funções de Grupo

```
tpch=> SELECT n.n_name, COUNT(c.c_mktsegment)
FROM customer c, nation n;
```


Consultas Ilegais - Funções de Grupo

```
tpch=> SELECT n.n_name, COUNT(c.c_mktsegment)  
FROM customer c, nation n
```

```
ERRO: coluna "n.n_name" deve aparecer na cláusula GROUP BY ou ser utilizada em uma função de agregação  
LINE 1: SELECT n.n_name, COUNT(c.c_mktsegment)
```

COLUNA AUSENTE

Consultas Ilegais - Funções de Grupo

```
tpch=> SELECT n.n_name, COUNT(c.c_mktsegment)
FROM customer c, nation n
WHERE AVG(c.c_acctbal)>1000
GROUP BY n.n_name;█
```


Consultas Ilegais - Funções de Grupo

```
tpch=> SELECT n.n_name, COUNT(c.c_mktsegment)
FROM customer c, nation n
WHERE AVG(c.c_acctbal)>1000
GROUP BY n.n_name;
```

```
ERRO: funções de agregação não são permitidas em WHERE
LINE 3: WHERE AVG(c.c_acctbal)>1000
```

WHERE NÃO É PERMITIDO COM
FUNÇÃO DE GRUPO

Funções de Grupo - HAVING

```
tpch=> SELECT c.c_mktsegment  
FROM customer c  
GROUP BY c.c_mktsegment  
HAVING AVG(c.c_acctbal)>4500;
```

```
c_mktsegment  
-----  
BUILDING  
HOUSEHOLD  
(2 rows)
```


Quando usar HAVING?

- É semelhante a WHERE porém utilizada somente com grupos.
- Pode ser utilizado ambas as cláusulas (WHERE e HAVING).

```
tpch=> SELECT c_mktsegment, AVG(c_acctbal)
FROM nation, customer
WHERE c_nationkey=n_nationkey and n_name='BRAZIL'
GROUP BY c_mktsegment
HAVING AVG(c_acctbal)>4400;
```


Quando usar HAVING?

```
tpch=> SELECT c_mktsegment, AVG(c_acctbal)
FROM nation, customer
WHERE c_nationkey=n_nationkey and n_name='BRAZIL'
GROUP BY c_mktsegment
HAVING AVG(c_acctbal)>4400;
```

c_mktsegment	avg
MACHINERY	4571.6550279776179057
AUTOMOBILE	4416.8059674657534247
BUILDING	4401.6134369885433715
HOUSEHOLD	4560.4621666666666667

(4 rows)

Exercícios

1. Crie uma consulta para exibir o tipo(p_type) e a média do preço de varejo dos itens de compra (part) em que o preço de varejo (p_retailprice) esteja entre R\$ 2095.10 e R\$ 2097.90, agrupados pelo tipo tipo . 7 linhas

p_type	avg
ECONOMY BRUSHED NICKEL	1481.2264988381099923
PROMO POLISHED COPPER	1504.8333435114503817
LARGE ANODIZED STEEL	1512.2303108415466262
SMALL POLISHED STEEL	1484.9336470588235294

Exercícios

2. Crie uma consulta para exibir a coluna `o_orderstatus`, o preço total (`o_totalprice`) máximo, agrupados pelo status das encomendas.

<code>o_orderstatus</code>	<code>max</code>
F	555285.16
O	530604.44
P	491549.57

(3 rows)

Exercícios

3. Crie uma consulta para exibir a média do custo de fornecimento (ps_supplycost) da tabela partsupp em que o tipo (p_type) da tabela part seja "PROMO POLISHED COPPER", agrupado pelo tipo;

```
      avg      |      p_type
-----+-----
497.8490820610687023 | PROMO POLISHED COPPER
(1 row)
```


Exercícios

4. Exibir o status do pedido (o_orderstatus) e o maior desconto (l_discount) com data entre 11 de novembro de 1998 e 31 de dezembro de 1998 agrupado pelo status do pedido.

o_orderstatus	max
0	0.10
(1 row)	

Exercícios

5. Crie uma consulta para exibir o nome (c_name) dos clientes e a média preço total (o_totalprice) dos clientes do BRAZIL agrupados pelo nome.

c_name	avg
Customer#000045317	165356.2316666666667
Customer#000045944	144617.2116666666667
Customer#000071138	195647.9370000000000
Customer#000099955	175615.7466666666667
Customer#000008591	77579.8133333333333
Customer#000079268	142927.944285714286
Customer#000021970	117280.7083333333333
Customer#000067054	173194.7346666666667
Customer#000123083	171362.5425000000000
Customer#000126881	128683.6860000000000

(4024 rows)