

Базы данных

08 | Множества группировок и поворот
данных

План

- Множества группировок
- ROLLUP и CUBE
- Идентификация группировок в результатах
- Поворот данных
- Использование PIVOT и UNPIVOT

Множества группировок

Синтаксис

- Раздел GROUPING SETS основан на предложении GROUP BY;
- Позволяем задавать множество группировок в одном запросе.

```
SELECT <список столбцов с агрегатами>  
FROM <источник>  
GROUP BY  
GROUPING SETS  
(  
    <column_name>, -- один или более столбцов  
    <column_name>, -- один или более столбцов  
    () -- круглые скобки, если надо агрегировать все строки  
);
```

Множества группировок

Пример

```
SELECT EmployeeID, CustomerID, SUM(Amount) AS TotalAmount
FROM Sales.SalesOrder
GROUP BY
GROUPING SETS(EmployeeID, CustomerID,());
```

	EmployeeID	CustomerID	TotalAmount
Итого для всех продаж {	<i>NULL</i>	<i>NULL</i>	256.23
	<i>NULL</i>	1	49.99
Подитог для каждого клиента {	<i>NULL</i>	2	107.49
	<i>NULL</i>	3	98.75
	1	<i>NULL</i>	107.49
Подитог для каждого сотрудника {	2	<i>NULL</i>	148.74

ROLLUP и CUBE

- ROLLUP обеспечивает множество группировок с комбинациями в иерархическом стиле:

```
SELECT StateProvince, City, COUNT(CustomerID) AS Customers  
FROM Sales.vCustomerDetails  
GROUP BY ROLLUP(StateProvince, City)  
ORDER BY StateProvince, City;
```

- CUBE обеспечивает множество группировок с перебором всех возможных комбинаций для группировки:

```
SELECT SalesPersonName, CustomerName, SUM(Amount) AS TotalAmount  
FROM Sales.vSalesOrders  
GROUP BY CUBE(SalesPersonName, CustomerName)  
ORDER BY SalesPersonName, CustomerName;
```

Идентификация группировок в результатах

- Несколько группировок затрудняют идентификацию источника каждой строки в результирующем наборе;
- NULL может появиться из данных или представлять значение-заполнитель в группировке;
- Через функцию GROUPING_ID можно определить группировку по столбцу (результат 1 или 0):

```
SELECT GROUPING_ID(SalesPersonName) AS SalesPersonGroup,  
       GROUPING_ID(CustomerName) AS CustomerGroup,  
       SalesPersonName, CustomerName, SUM(Amount) AS TotalAmount  
FROM Sales.vSalesOrders  
GROUP BY CUBE(SalesPersonName, CustomerName)  
ORDER BY SalesPersonName, CustomerName;
```

GROUPING_ID для нескольких столбцов

```
SELECT a, b, c, SUM(d), GROUPING_ID(a,b,c)
FROM T GROUP BY <список group by>
```

Статистически обработанные столбцы	Входные данные GROUPING_ID (a, b, c) = GROUPING(a) + GROUPING(b) + GROUPING(c)	Выходные данные GROUPING_ID ()
a	100	4
b	010	2
c	001	1
ab	110	6
ac	101	5
bc	011	3
abc	111	7

GROUPING_ID () возвращает битовую карту типа integer, в которой N самых младших битов могут быть установлены (1).

Установленный бит означает, что соответствующий аргумент не является столбцом группирования для указанной выходной строки.

Применительно к одиночному запросу группировки функция GROUPING (<выражение-столбец>) эквивалентна GROUPING_ID (<выражение-столбец>), и обе эти функции возвращают 0.

DEMO

Множества группировок

Поворот данных (PIVOT)

- Поворот данных – это транспонирование данных из построчной ориентации в ориентацию на основе столбцов;
- Уникальные значения из одного столбца проецируются в заголовки столбцов — может происходить агрегация:

OrderID	ProductID	Category	Revenue
1023	1	Bikes	1078.75
1023	15	Accessories	52.00
1023	21	Accessories	124.90
1024	1	Bikes	2491.00
1025	3	Bikes	1067.49
1025	15	Accessories	125.99
1025	35	Clothing	26.57
1025	36	Clothing	5.78

OrderID	Bikes	Accessories	Clothing
1023	1078.75	176.90	<i>NULL</i>
1024	2491.00	<i>NULL</i>	<i>NULL</i>
1025	1067.49	125.99	32.35



```
SELECT OrderID, Bikes, Accessories, Clothing
FROM
    (SELECT OrderID, Category, Revenue FROM Sales.SalesDetails) AS sales
PIVOT (SUM(Revenue) FOR Category IN([Bikes], [Accessories], [Clothing])) AS pvt
```

Обратный поворот данных (UNPIVOT)

- Обратный поворот – это поворот данных из ориентации на основе столбцов в ориентацию на основе строк;
- Распространяет или разделяет значения из одной строки-источника в одну или несколько целевых строк;
- Каждая строка-источник становится одной или несколькими строками в результате, основанном на количестве поворотных столбцов:

OrderID	Bikes	Accessories	Clothing
1023	1078.75	176.90	NULL
1024	2491.00	NULL	NULL
1025	1067.49	125.99	32.35

```
SELECT OrderID, Category, Revenue
FROM
    (SELECT OrderID, Bikes, Accessories, Clothing FROM Sales.SalesByCat) AS pvt
UNPIVOT (Revenue FOR Category IN([Bikes], [Accessories], [Clothing])) AS unpvt
```

OrderID	Category	Revenue
1023	Bikes	1078.75
1023	Accessories	176.90
1024	Bikes	2491.00
1025	Bikes	1067.49
1025	Accessories	125.99
1025	Clothing	32.35

DEMO

Поворот и обратный поворот данных

Изучено

- Множества группировок
- ROLLUP и CUBE
- Идентификация группировок в результатах
- Поворот данных
- Использование PIVOT и UNPIVOT