

SQL 必知必会 9 - 汇总数据

SQL 必知必会 9 - 汇总数据

9.1 聚合函数

9.1.1 AVG() 函数

9.1.2 COUNT() 函数

9.1.3 MAX() 函数

9.1.4 MIN() 函数

9.1.5 SUM() 函数

9.2 聚合不同值

9.3 组合聚合函数

9.4 小结

9.1 聚合函数

聚合函数 **aggregate function**

对某些行运行的函数，计算并返回一个值。

表9-1 SQL聚集函数

函 数	说 明
AVG()	返回某列的平均值
COUNT()	返回某列的行数
MAX()	返回某列的最大值
MIN()	返回某列的最小值
SUM()	返回某列值之和

9.1.1 **AVG()** 函数

返回平均值。

返回 **Products** 表中所有产品的平均价格：

```
SELECT AVG(prod_price) AS avg_price
```

```
FROM Products;
```

返回特定供应商所提供产品的平均价格：

```
SELECT AVG(prod_price) AS avg_price
FROM Products
WHERE vend_id = 'DLL01';
```

说明

- `AVG()` 只用于单个列。为了获得多个列的平均值，必须使用多个 `AVG()` 函数。
- `AVG()` 函数忽略列值为 `NULL` 的行。

9.1.2 `COUNT()` 函数

`COUNT()` 函数进行计数。

两种使用方式：

- 使用 `COUNT(*)` 对表中行的数目进行计数，不管表列中包含的是空值 `NULL` 还是非空值。
- 使用 `COUNT(column)` 对特定列中具有值的行进行计数，忽略 `NULL` 值。

返回 `Customers` 表中顾客的总数：

```
SELECT COUNT(*) AS num_cust
FROM Customers;
```

只对具有电子邮件地址的客户计数：

```
SELECT COUNT(cust_email) AS num_cust
FROM Customers;
```

说明

`NULL` 值

- 如果指定列名，则 `COUNT()` 函数会忽略指定列的值为空的行
- 如果 `COUNT()` 函数中用的是星号 `*`，则不忽略。

9.1.3 `MAX()` 函数

返回指定列中的最大值。

`MAX()` 要求指定列名。

```
SELECT MAX(prod_price) AS max_price
FROM Products;
```

说明

- 许多（并非所有）`DBMS` 允许将它用来返回任意列中的最大值，包括返回文本列中的最大值。
在用于文本数据时，`MAX()` 返回按该列排序后的最后一行。
- `MAX()` 函数忽略列值为 `NULL` 的行。

9.1.4 `MIN()` 函数

返回指定列的最小值。

`MIN()` 要求指定列名。

```
SELECT MIN(prod_price) AS min_price
FROM Products;
```

- 大多数 `DBMS` 可以对非数值数据使用 `MIN()`。在用于文本数据时，`MIN()` 返回该列排序后最前面的行。
- `MIN()` 函数忽略列值为 `NULL` 的行。

9.1.5 `SUM()` 函数

`SUM()` 用来返回指定列值的和（总计）。

查询订单号为 `20005` 的订单总数量。

```
SELECT SUM(quantity) AS items_ordered
FROM OrderItems
WHERE order_num = 20005;
```

查询订单号为 `20005` 的订单总金额。

```
SELECT SUM(item_price*quantity) AS total_price
FROM OderItems
WHERE order_num = 20005;
```

说明

所有聚合函数都可用来执行多个列上的计算。

`SUM()` 函数忽略列值为 `NULL` 的行

9.2 聚合不同值

五个聚合函数的口诀：均数大小和

五个聚合函数都可以指定 `ALL` 和 `DISTINCT`：

- 对所有行执行计算，指定 `ALL` 参数或不指定参数（默认就是 `ALL`）
- 只包含不同的值，指定 `DISTINCT` 参数

`Microsoft Access` 在聚合函数中不支持 `DISTINCT`

要在 `Access` 得到类似的结果，需要使用子查询把 `DISTINCT` 数据返回到外部 `SELECT COUNT(*)` 语句

```
SELECT AVG(DISTINCT prod_price) AS avg_price
FROM Products
WHERE vend_id = 'DLL01';
```

`DISTINCT` 不能用于 `COUNT(*)`

虽然 `DISTINCT` 可用于 `MIN()` 和 `MAX()`，但这样做实际上没有价值。

有的 `DBMS` 还支持其他参数，如支持对查询结果的子集进行计算的 `TOP` 和 `TOP PERCENT`。

9.3 组合聚合函数

`SELECT` 语句可根据需要包含多个聚合函数。

```
SELECT COUNT(*) AS num_items,
       MIN(prod_price) AS price_min,
       MAX(prod_price) AS price_max,
       AVG(prod_price) AS price_avg
FROM Products;
```

注意取别名不应该使用表中的实际列名。

虽然这样做也算合法，但许多 SQL 实现不支持，可能会产生模糊的错误消息。

9.4 小结