

文章

[姚鑫](#) · 十月 18, 2021 阅读大约需 8 分钟

第四十九章 SQL命令 GROUP BY

第四十九章 SQL命令 GROUP BY

SELECT子句，它根据一个或多个列对查询的结果行进行分组。

大纲

```
SELECT ...  
GROUP BY field {,field2}
```

参数

- field - 从其中检索数据的一个或多个字段。
单个字段名或以逗号分隔的字段名列表。

描述

GROUP BY是SELECT命令的一个子句。

可选的GROUP BY子句出现在FROM子句和可选的WHERE子句之后，可选的HAVING和ORDER BY子句之前。

GROUP BY子句接受查询的结果行，并根据一个或多个数据库列将它们分成单独的组。当将SELECT与GROUP BY结合使用时，将为GROUP BY字段的每个不同值检索一行。GROUP BY将NULL(没有指定值)字段作为一个独立的值组。

GROUP BY子句在概念上类似于IRIS聚合函数扩展关键字%FOREACH，但是GROUP BY操作整个查询，而%FOREACH允许在子填充上选择聚合，而不限制整个查询填充。

GROUP BY可以在INSERT命令的SELECT子句中使用。
不能在UPDATE或DELETE命令中使用GROUP BY。

指定字段

GROUP BY子句最简单的形式指定单个字段，如GROUP BY City。
这将为每个唯一的City值选择任意一行。
还可以指定以逗号分隔的字段列表，将其组合值视为单个分组术语。
它为每个City和Age值的唯一组合选择任意一行。
因此，GROUP BY City, Age返回与GROUP BY Age, City相同的结果。

字段必须通过列名指定。

有效的字段值包括以下内容:列名(GROUP BY City);

%ID(返回所有行);

指定列名的标量函数(GROUP BY ROUND(Age, -1));

指定列名的排序规则函数(GROUP BY %EXACT(City))。

不能通过列别名指定字段;

尝试这样做会产生SQLCODE -29错误。

不能通过列号指定字段;

这被解释为一个文字并返回一行。

不能指定聚合字段;

尝试这样做将生成SQLCODE -19错误。

不能指定子查询;

这被解释为一个文字并返回一行。

GROUP BY StreamField操作流字段的OID，而不是它的实际数据。

因为所有流字段oid都是唯一的值，GROUP BY对实际的流字段重复数据值没有影响。

GROUP BY StreamField将流字段为NULL的记录数量减少为一条记录。

GROUP BY子句可以使用箭头语法(->)操作符在非基表的表中指定字段。

例如:GROUP BY Company->Name。

在GROUP BY子句中指定一个字面值作为字段值返回1行;

返回哪一行是不确定的。

因此，指定7、'Chicago'、"、0或NULL都返回1行。

但是，如果在逗号分隔的列表中指定一个字面值作为字段值，则该字面值将被忽略，并且GROUP BY将为指定字段名的每个惟一组合选择任意一行。

具有GROUP BY和DISTINCT BY的聚合函数

在计算聚合函数之前应用GROUP BY子句。

在下面的示例中，COUNT聚合函数计算每个GROUP BY组中的行数:

```
SELECT Home_State,COUNT(Home_State)
FROM Sample.Person
GROUP BY Home_State
```

在计算聚合函数之后应用DISTINCT BY子句。

在下面的例子中，COUNT聚合函数计算整个表中的行数:

```
SELECT DISTINCT BY(Home_State) Home_State,COUNT(Home_State)
FROM Sample.Person
```

为了计算整个表的聚合函数，而不是GROUP BY组，可以指定一个选择项子查询:

```
SELECT Home_State,(SELECT COUNT(Home_State) FROM Sample.Person)
FROM Sample.Person
GROUP BY Home_State
```

当选择列表由聚合字段组成时，不应将GROUP BY子句与DISTINCT子句一起使用。

例如，下面的查询旨在返回共享相同HomeState的不同数量的人:

```
/* ??????DISTINCT??? */
/* ?????????????? */
SELECT DISTINCT COUNT(*) AS mynum
FROM Sample.Person
GROUP BY Home_State
ORDER BY mynum
```

这个查询没有返回预期的结果，因为它没有应用DISTINCT关键字。
要同时应用DISTINCT聚合和GROUP BY子句，请使用子查询，如下例所示：

```
SELECT DISTINCT *
FROM (SELECT COUNT(*) AS mynum
      FROM Sample.Person
      GROUP BY Home_State) AS Sub
ORDER BY Sub.mynum
```

此示例成功返回共享相同HomeState的不同人数。
例如，如果任何HomeState被8个人共享，查询返回8。

如果查询仅由聚合函数组成且不返回表中的任何数据，则返回%ROWCOUNT=1，并为聚合函数返回一个空字符串(或0)值。

例如：

```
SELECT AVG(Age) FROM Sample.Person WHERE Name %STARTSWITH 'ZZZZ'
```

但是，如果这种类型的查询包含GROUP BY子句，它将返回%ROWCOUNT=0，并且聚合函数值仍未定义。

飘絮，字母大小写和优化

本节描述GROUP BY如何处理只有字母大小写不同的数据值。

- 组合字母变体在一起(返回大写字母):

默认情况下，GROUP BY根据创建字段时为其指定的排序规则将字符串值分组。

IRIS有一个默认的字符串排序规则，可以为每个名称空间设置；
所有名称空间的初始字符串排序规则默认值是SQLUPPER。
因此，除非另有说明，通常GROUP BY排序规则不区分大小写。

GROUP BY根据字段的大写字母排序规则，使用SQLUPPER排序规则对字段的值进行分组。

只有字母大小写不同的字段值被分组在一起。

分组字段值全部以大写字母返回。

这样做的性能优势在于允许GROUP BY为字段使用索引，而不是访问实际的字段值。

因此，只有在一个或多个选定字段的索引存在时才有意义。

它的结果是group by字段值全部以大写字母返回，即使实际数据值中没有一个都是大写字母。

- 组合字母大小写变体在一起(返回实际的字母大小写):

GROUP BY可以将字母大小写不同的值分组在一起，并使用实际的字段字母大小写值返回分组的字段值(随机选择)。

这样做的好处是返回的值是实际值，显示数据中至少一个值的字母大小写。

它的性能缺点是不能使用字段的索引。

可以通过对select-item字段应用%EXACT排序函数来为单个查询指定这个值。

- 不要将不同的字母组合在一起(返回实际的字母):

通过对GROUP BY字段应用%EXACT排序功能，GROUP BY可以对值进行区分大小写的分组。

这样做的好处是将每个字母变体作为一个单独的组返回。


```
&sql(CLOSE EmpCursor11)
}
```

```
DHC-APP>d ##class(PHA.TEST.SQLCommand).GroupBy()
```

```
RowID: 999 row count: 1 Name=O'Rielly,Chris H. State=MS
RowID: 999 row count: 2 Name=Orwell,John V. State=MT
RowID: 999 row count: 3 Name=Zevon,Heloisa O. State=MI
RowID: 999 row count: 4 Name=Kratzmann,Emily Z. State=MO
RowID: 999 row count: 5 Name=Hanson,George C. State=MD
RowID: 999 row count: 6 Name=Zucherro,Olga H. State=MN
RowID: 999 row count: 7 Name=Gallant,Thelma Q. State=MA
RowID: 999 row count: 8 Name=Xiang,Kirsten U. State=ME
```

查询行为的这种更改仅适用于基于游标的嵌入式SQL SELECT查询。动态SQL SELECT查询和非游标嵌入式SQL SELECT查询从未设置%ROWID。

事务提交的更改

包含GROUP BY子句的查询不支持READ COMMITTED隔离级别。在定义为READ COMMITTED的事务中，不带GROUP

BY子句的SELECT语句仅返回已提交的数据修改；换句话说，它返回当前事务之前的数据状态。带有GROUP BY子句的SELECT语句返回所做的所有数据修改，无论它们是否已提交。

示例

下面的示例按名称的首字母对名称进行分组。它返回首字母、共享该首字母的姓名计数以及一个Name值的示例。名称使用其SQLUPPER排序规则进行分组，而不考虑实际值的字母大小写。请注意，名称SELECT-ITEM包含大写首字母；%Exact排序规则用于显示实际的Name值：

```
SELECT Name AS Initial,COUNT(Name) AS SameInitial,%EXACT(Name) AS Example
FROM Sample.Person GROUP BY %SQLUPPER(Name,2)
```

[#SQL #Caché](#)

源

URL:

<https://cn.community.intersystems.com/post/%E7%AC%AC%E5%9B%9B%E5%8D%81%E4%B9%9D%E7%AB%A0-sql%E5%91%BD%E4%BB%A4-group>