# 概述

正确的用法：

select distinct column1,column2 from table;

这样就可以查询出来column1和column2不重复的行。

错误的用法：

select column1,distinct column2 from table;

这样是错误用法，distinct只能用在最左边。

# 原理

## Sort unique

在10gR2以前，distinct走的是Sort unique，group by走的是Sort group by，两者都需要进行排序；10gR2以后，默认情况下distinct改走Hash unique，group by走Hash group by。

## Hash unique

distinct oracle改进后使用hash unique算法，效率已经有了很大提高。但是仍然会对结果排序，所以性能有一定的影响，所以在非必要的情况下不建议使用。

# 优化

https://blog.csdn.net/shengsummer/article/details/31010219

## group by

使用group by代替distinct：

<http://www.itpub.net/thread-1392256-1-1.html>

## exists

使用exists代替distinct：

<http://blog.itpub.net/10856805/viewspace-1000690>

## 调整表扫描顺序

在多个表去重操作时，可以根据表的大小调整去重操作的顺序。

比如，存在表t1，t2，t3，大小关系为t1>t2>t3，操作SQL为：select dustinct t1.\*,t2,\*,t3,\* from t1,t2,t3;，按照原来的逻辑去重操作顺序为：t1，t2，t3，但是这样效率较低，调整去重操作顺序为：t3，t2，t1，这样t3与t2去重后数据变得更少了。