**数据库实验报告**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **实 验 人：** | **凌国明** | | **学 号：** | **21307077** | **日 期：** | **2023年9月14日** |
| **院（系）：** | **计算机学院** | | | **专业（班级）：** | **计算机科学与技术** | |
| **联系方式：** |  | | |  |  | |
| **实验题目：** | | **1.3 数据高级查询** | | | | |

* + - 1. **实验目的**

掌握 SQL 嵌套查询和集合查询等各种高级查询的设计方法。

* + - 1. **实验内容和要求**

针对 TPC-H 数据库，正确分析用户查询要求，设计各种嵌套查询和集合查询。

* + - 1. **实验重点和难点**

实验重点：嵌套查询

实验难点：相关子查询、多层 exist 嵌套查询

* + - 1. **实验工具**

MySQL、SQL Server、Navicat

* + - 1. **实验过程**

（1）In 嵌套查询

查询订购了“佳能”制造的“佳能墨盒”的顾客。

方法一

执行以下SQL语句：

SELECT customer.custkey, customer.name

FROM orders

JOIN lineitem ON orders.orderkey=lineitem.orderkey

JOIN customer ON orders.custkey=customer.custkey

WHERE lineitem.partkey in (SELECT partkey FROM part WHERE name='佳能墨盒')

部分结果截屏如下：



方法二

执行以下SQL语句：

SELECT customer.custkey, customer.name

FROM orders

JOIN lineitem ON orders.orderkey = lineitem.orderkey

JOIN customer ON orders.custkey = customer.custkey

JOIN part on lineitem.partkey = part.partkey

WHERE part.name = '佳能墨盒'

部分结果截屏如下：



（2）单层 exists 嵌套查询

查询没有订购过“佳能”制造的“佳能墨盒”的顾客。

执行以下SQL语句：

SELECT c.custkey, c.name

FROM customer AS c

WHERE NOT EXISTS(

SELECT 0

FROM orders o

JOIN lineitem l ON o.orderkey = l.orderkey

JOIN part p on l.partkey = p.partkey

WHERE p.name = '佳能墨盒' AND c.custkey = o.custkey

);

部分结果截屏如下：



（3）双层 exists 嵌套查询（注：由于表太大，可能非常耗时！）

查询至少购买过顾客“齐光远”购买过的全部零件的顾客姓名。

执行以下SQL语句：

SELECT DISTINCT c.name

FROM customer c

WHERE EXISTS (

SELECT 1

FROM orders o1

JOIN lineitem l1 ON o1.orderkey = l1.orderkey

JOIN part p1 ON l1.partkey = p1.partkey

WHERE c.custkey = o1.custkey AND p1.partkey IN (

SELECT p.partkey

FROM orders o

JOIN lineitem l ON o.orderkey = l.orderkey

JOIN part p ON l.partkey = p.partkey

JOIN customer cc ON o.custkey = cc.custkey

WHERE cc.`name` = '齐光远'

)

) AND NOT EXISTS (

SELECT 1

FROM part p2

WHERE p2.partkey IN (

SELECT p.partkey

FROM orders o

JOIN lineitem l ON o.orderkey = l.orderkey

JOIN part p ON l.partkey = p.partkey

JOIN customer cc ON o.custkey = cc.custkey

WHERE cc.`name` = '齐光远'

)

AND p2.partkey NOT IN(

SELECT p.partkey

FROM orders o

JOIN lineitem l ON o.orderkey = l.orderkey

JOIN part p ON l.partkey = p.partkey

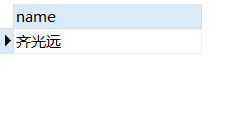
JOIN customer cc ON o.custkey = cc.custkey

WHERE cc.`name` = c.`name`

)

);

部分结果截屏如下：



**注意，这里其实只用第二部分not exists就能实现题意的功能，但是加上第一部分的exists判断，能大量缩短查询时间。（第一部分先判断这个顾客购买产品是否是“齐光远”买过的）**

（4）From 子句中的嵌套查询

查询订单平均金额超过 25 万元的顾客中的中国籍顾客信息。

执行以下SQL语句：

SELECT \*

FROM customer c

JOIN (

SELECT custkey

FROM orders

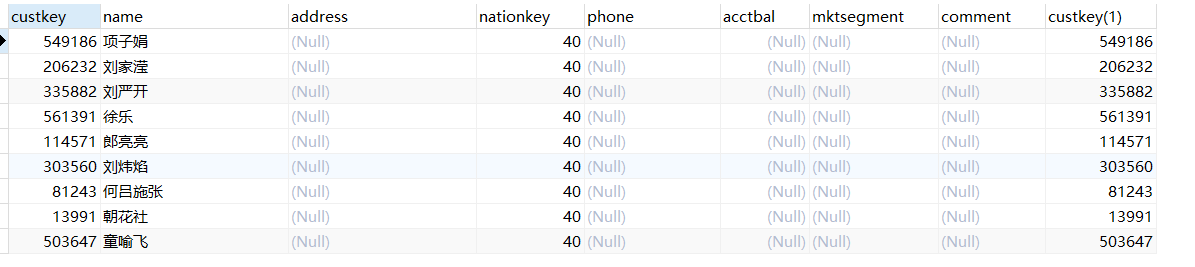
GROUP BY custkey

HAVING AVG(totalprice) > 250000

) subquery ON c.custkey = subquery.custkey

WHERE c.nationkey = (SELECT nationkey FROM nation WHERE `name`='中国')

部分结果截屏如下：



（5）集合查询（交）

查询顾客“陈朋”和“宋金保” 都订过的全部零件的信息。

执行以下SQL语句：

SELECT part.partkey, part.`name`

FROM part

WHERE part.partkey in (

SELECT p.partkey

FROM orders o

JOIN lineitem l ON o.orderkey = l.orderkey

JOIN part p ON l.partkey = p.partkey

JOIN customer c ON o.custkey = c.custkey

WHERE c.`name` = '陈朋'

)

AND part.partkey in (

SELECT p1.partkey

FROM orders o1

JOIN lineitem l1 ON o1.orderkey = l1.orderkey

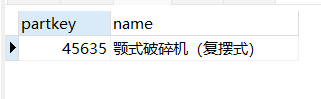
JOIN part p1 ON l1.partkey = p1.partkey

JOIN customer c1 ON o1.custkey = c1.custkey

WHERE c1.`name` = '宋金保'

)

部分结果截屏如下：



（6）集合查询（并）

查询顾客“陈朋”和“宋金保”订购的全部零件的信息。

执行以下SQL语句：

SELECT p.partkey, p.`name`

FROM orders o

JOIN lineitem l ON o.orderkey = l.orderkey

JOIN part p ON l.partkey = p.partkey

JOIN customer c ON o.custkey = c.custkey

WHERE c.`name` = '陈朋'

UNION

SELECT p1.partkey, p1.`name`

FROM orders o1

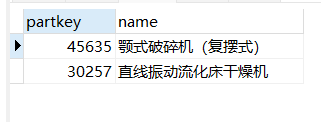
JOIN lineitem l1 ON o1.orderkey = l1.orderkey

JOIN part p1 ON l1.partkey = p1.partkey

JOIN customer c1 ON o1.custkey = c1.custkey

WHERE c1.`name` = '宋金保'

部分结果截屏如下：



（7） 集合查询（差）

查询顾客“宋金保”订过而顾客“陈朋”没订过的零件的信息。

执行以下SQL语句：

SELECT part.partkey, part.`name`

FROM part

WHERE part.partkey in (

SELECT p.partkey

FROM orders o

JOIN lineitem l ON o.orderkey = l.orderkey

JOIN part p ON l.partkey = p.partkey

JOIN customer c ON o.custkey = c.custkey

WHERE c.`name` = '宋金保'

)

AND part.partkey NOT in (

SELECT p1.partkey

FROM orders o1

JOIN lineitem l1 ON o1.orderkey = l1.orderkey

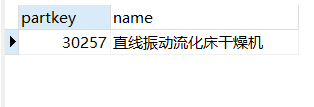
JOIN part p1 ON l1.partkey = p1.partkey

JOIN customer c1 ON o1.custkey = c1.custkey

WHERE c1.`name` = '陈朋'

)

部分结果截屏如下：



* + - 1. **与实验结果相关的文件**

无

* + - 1. **实验总结**

在这次实验中，我学习了IN嵌套查询、FROM嵌套子查询以及EXISTS单层和双层嵌套查询，这些都是数据库查询中的重要概念。其中，双层EXISTS嵌套查询成为了我的挑战之一。由于查询的表很大，运行时间较长，这促使我寻求优化的方法。我通过对查询的改进（见上方相应部分），使查询能在瞬间完成，这让我很有成就感。这个过程教会了我不仅要理解查询语法，还要深入思考如何提高查询效率，这对于将来在实际项目中处理大型数据集至关重要。

此外，我也学到了关于集合操作的知识，包括交、并、差等。尽管SQL没有提供现成的关键词来执行这些操作，但通过编程的方式去实现它们让我更好地理解了数据处理的灵活性。在实际应用中，这些集合操作可以帮助我们处理和分析复杂的数据，对于数据库查询和数据处理的技能提升至关重要。