DESK报表开发基础测试题

1. **基础知识（2小题，每题5分，合计10分）**
2. 请写出至少7个你所知道的DB2的数据类型；简述char和varchar的区别。**（7分）**
3. 请写出至少3个你所知道的unix的一些查看文件内容的命令；简述文件存取权限的**（3分）**
4. **考察表关联（3小题，合计25分）**

数据模型如下：

表T1的数据结构:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 字段英文名 | 字段中文名 | 类型 | 主键标志 | 注释 |
| Rec\_no | 记录号 | CHAR(3) | Y |  |
| Ci\_no | 客户号 | CHAR(6) | N |  |
| Cust\_Type | 客户类型 | CHAR(2) | N |  |
| Cre\_dt | 开户日期 | Date | N |  |
| Cus\_sts | 客户状态 | Char(1) | N | Y—正常  N—无效 |

表T1的数据

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Rec\_no | Ci\_no | Cust\_type | Cre\_dt | Cus\_sts |
| 123 | 111111 | 01 | 2010-11-15 | Y |
| 234 | 222222 | 02 | 2011-09-01 | N |
| 345 | 333333 | 02 | 2012-01-09 | Y |
| 456 | 444444 | 01 | 2012-09-08 | Y |

表T2的数据结构：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 字段英文名 | 字段中文名 | 类型 | 主键标志 | 注释 |
| Ci\_no | 客户号 | CHAR(6) | Y |  |
| Ac\_no | 客户账号 | CHAR(9) | Y |  |
| Bal | 账号余额 | DECIMAL(15,2) | N |  |

表T2的数据:

|  |  |  |
| --- | --- | --- |
| Ci\_no char(6) | Ac\_no  char(9) | Bal decimal(15,2) |
| 222222 | 123456789 | 1000.28 |
| 333333 | 123454321 | 5000.00 |

**问题1**.请写出

select t1.rec\_no,t1.ci\_no, coalesce(t2.ac\_no,'888888888') as ac\_no,round(coalesce(t2.bal,0.00) ,1 ) as bal from T1 left join T2 on T1.ci\_no=T2.ci\_no

的结果，填于下表中。**（8分）**

|  |  |  |  |
| --- | --- | --- | --- |
| REC\_NO | CI\_NO | AC\_NO | BAL |
| *123* | *111111* | 888888888 | *0.0* |
| *234* | *222222* | 123456789 | 1000.3 |
| *345* | *333333* | 123454321 | *5000.0* |
| *456* | *444444* | 888888888 | *0.0* |
| … | … | … | … |

**问题2**： 请编写sql统计在9月份开户且账户余额不为0的有效客户数。**（8分）**

SELECT COUNT(DISTINCT ci\_no) FROM Customers2 c JOIN Accounts2 a ON c.ci\_no=a.ci\_no

WHERE to\_char(cre\_dt,'mm')='09' AND a.bal>0 AND cus\_sts='Y';

**问题3**： 如果在上一题中，遇到性能问题（当数据量非常大时，尤其是T2表的数据量远超T1表时），请从index入手，分析如何可以提高查询统计的性能; 在关联表的顺序上，请问上一题的sql 写法是否合理，是否需要调整顺序，给出理由 ：**（9分）**

1. **考察数据展现（1小题，合计8分）**

表TT的数据模型如下：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 字段英文名 | 字段中文名 | 类型 | 主键标志 | 注释 |
| Year | 年度 | Int | N |  |
| Quarter | 季度 | Int | N |  |
| Amt | 金额 | Int | N |  |

目前的数据如下：

YEAR QUARTER AMT

----------- ----------- -----------------

2011 1 110

2011 2 120

2011 3 130

2011 4 140

2012 4 240

2012 3 230

2012 2 220

2012 1 210

**问题1**：请编写sql，按照以下格式显示数据：（10分）

YEAR Q1 Q2 Q3 Q4

----------- ----------------- ----------------- ----------------- -----------------

2011 110 120 130 140

2012 210 220 230 240

SELECT \* FROM tt PIVOT(MAX(amt) FOR quarter IN(1 Q1,2 Q2,3 Q3,4 Q4));

1. **拉链算法理解（2小题，合计20分）**

拉链算法，顾名思义，有着连接、衔接的意思.拉链算法是目前数据仓库领域比较好的算法之一..通用非常广.记录数据量很大且为全量实体记录历史的操作。

T3：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 字段英文名 | 字段中文名 | 类型 | 主键标志 | 注释 |
| Ac\_no | 账号 | CHAR(9) | N |  |
| Bal | 账号余额 | Decimal(15,2) | N |  |
| St\_dt | 拉链开始日期 | Date | N |  |
| End\_dt | 拉链结束日期 | Date | N |  |

假设在最初(2012-01-05)，有以下数据：

|  |  |  |  |
| --- | --- | --- | --- |
| Ac\_no | Bal | St\_dt | End\_dt |
| 123456789 | 100 | 2012-01-05 | 2999-12-31 |
| 123454321 | 200 | 2012-01-05 | 2999-12-31 |

当在（2012-01-06），账户情况发生了变化：

(1) Ac\_no为123456789 的余额（Bal）变成500

(2) 新增了账号888812345，其余额（Bal）为1000。

**问题1**：请写出T3表在2012-01-06批量后的数据情况：**（12分）**

|  |  |  |  |
| --- | --- | --- | --- |
| Ac\_no | Bal | St\_dt | End\_dt |
| 123456789 | 100 | 2012-01-05 | 2012-01-05 |
| 123454321 | 200 | 2012-01-05 | 2999-12-31 |
| 123456789 | 500 | 2012-01-06 | 2999-12-31 |
| 888812345 | *1000* | 2012-01-06 | 2999-12-31 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**问题2**：假设2012-05-10是一个普通的会计日期，请使用sql 列出当前余额大于1000的记录。**（8分）**

Select \* from t3 where st\_dt<=to\_date(‘2012-05-10’,’yyyy-mm-dd’) and end\_dt>=to\_date(‘2012-05-10’,’yyyy-mm-dd’) and bal>1000;

1. **SHELL程序编写（1小题，合计12分）**

AIX（unix）下，有一存放数据的目录，数据以日期为文件夹进行存放，如：

20120101

20120102

20120103

。。。

请编写shell程序，对该目录下的所有日期里面的，以”.dat” 结尾的文件进行压缩。(请务必把重点使用的方法写明)

方式一：

#!/bin/bash

cd /root

fs=`find www -name \*.bat`

#压缩打包

tar -zcvf bats.tar.gz $fs

方式二：

#!/bin/bash

#获取文件夹列表

dirs=`ls /root/www`

i=0

for d in $dirs

do

#判断是否是一个目录

if [ -d $d ];then

#获取目录下面所有以。bat结尾的文件

fs=`ls $d | grep '\.bat$'`

# echo $fs

for f in $fs

do

i=$((i+1))

if [ $i -eq 1 ];then

tar -c -f bats.tar $d/$f

else

tar -r -f bats.tar $d/$f

fi

done

fi

done;

#进行文件的压缩

gzip bats.tar

**六．案例分析：客户产品交叉销售（3小题，合计25分）**

数据模型如下：

T4: (客户持有产品汇总表)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 字段英文名 | 字段中文名 | 类型 | 主键标志 | 注释 |
| Rec\_No | 序列号 | Char(9) | Y |  |
| Cust\_Id | 客户编号 | CHAR(13) | Y |  |
| Product\_Id | 产品编号 | CHAR(8) | Y |  |
| Amt | 产品余额 | Decimal(15,2) | N |  |
| Txn\_Dt | 交易日期 | DATE | N |  |
| Org\_Id | 分行编号 | CHAR(6) | N |  |

**问题1**：编写一段SQL，统计2012年9月深圳分行（310111）各产品的客户数以及余额汇总，并按照产品客户数降序排列**（8分）**

|  |  |  |
| --- | --- | --- |
| 产品编号 | 客户数 | 余额汇总 |
| Xxxxxx | 123456 | 88888888 |
| Xxxxxx | 100000 | 99999999 |
| .. | .. | .. |

*Select product\_id 产品编号,count(cust\_id)客户数,sum(amt) 余额汇总from t4 where txn\_dt between to\_date(‘2012-9-1’,’yyyy-mm-dd’) and to\_date(‘2012-9-30’,’yyyy-mm-dd’) and org\_id=’*310111*’ group by product\_id order by 客户数 desc;*

**问题2**: 编写一段SQL，找到2012年9月份全行持有产品”05108100”的客户列表中持有其他产品的客户数，并按照产品客户数降序排列**（9分）**

|  |  |
| --- | --- |
| 产品编号 | 客户数 |
| Xxxxxx | 123456 |
| Xxxxxx | 100000 |
| .. | .. |

方式一：

SELECT product\_id 产品编号,COUNT(\*) 客户数 FROM t4 tt WHERE EXISTS(

SELECT cust\_id,COUNT(\*) FROM t4 t WHERE EXISTS (

SELECT \* FROM t4 WHERE product\_id='p03' AND t.cust\_id=cust\_id )

GROUP BY cust\_id HAVING COUNT(\*)>1 AND tt.cust\_id=cust\_id)

GROUP BY product\_id ORDER BY 客户数 desc;

方式二：

SELECT product\_id 产品编号,COUNT(\*) 客户数 FROM t4 WHERE EXISTS (

SELECT cust\_id FROM t4 t WHERE EXISTS (

SELECT \* FROM t4 WHERE product\_id='p03' AND t.cust\_id=cust\_id)

AND product\_id<>'p03' AND t4.cust\_id=cust\_id) GROUP BY product\_id ORDER BY 客户数 desc;

**问题3**：编写一段SQL，找到2012年9月份全行所有产品余额总计超过1000万的前10家分行，并按照产品余额总计降序排列**（8分）**

|  |  |  |
| --- | --- | --- |
| 分行编号 | 产品编号 | 余额总计 |
| Xxxxxx | Xxxxxx | 12345600 |
| Xxxxxx | Xxxxxx | 10000000 |
|  | .. | .. |

SELECT \* FROM(

SELECT org\_id,product\_id,余额,总余额,dense\_rank()OVER(ORDER BY 总余额 DESC) 排名 FROM(

SELECT a.\*,SUM(余额)over(PARTITION BY org\_id) 总余额 FROM(

SELECT org\_id,product\_id,SUM(amt) 余额 FROM t4 GROUP BY org\_id,product\_id) a) WHERE 总余额>1000)

WHERE 排名<=10;