京东大数据专家认证培训SQL基础知识

**SQL定义：**SQL是一种面向数据库的通用数据处理语言规范，能完成以下几类功能：提取查询数据，插入修改删除数据，生成修改和删除数据库对象，数据库安全控制，数据库完整性及数据保护控制。  
  
**SQL分类：**  
DDL—数据定义语言（CREATE，ALTER，DROP，DECLARE）  
DML—数据操纵语言（SELECT，DELETE，UPDATE，INSERT）  
DCL—数据控制语言（GRANT，REVOKE，COMMIT，ROLLBACK）  
  
**db2数据类型**  
CHAR()：定长字符串  最大长度为 254     
VARCHAR()：变长字符  最大长度为 4000    
SMALLINT：短整型数字  长度为 2 字节   
INTEGER：整型数字  长度为 4 字节   
REAL：单精度浮点  32 位近似值   
DOUBLE：双精度浮点  64 位近似值  
DECIMAL(m,n)：数字  精度为m小数位为n  
DATE：日期时间  
TIME：日期时间  
TIMESTAMP：日期时间  
  
DDL—数据库定义语言：直接提交的。  
   CREATE：用于创建数据库对象。  
   DECLARE：除了是创建只在过程中使用的临时表外，DECLARE语句和CREATE语句非常相似。唯一可以被声明的对象是表。并且必须放入用户临时表空间。    
   DROP：可以删除任何用CREATE（数据库对象）和DECLARE（表）创建的对象。  
   ALTER：允许修改某些数据库对象的信息。不能修改索引。  
下面主要基于对象介绍基本的语法：  
1、数据库：  
  创建数据库：CREATE DATABASE database-name [USING CODESET codeset TERRITORY territory]   
  注：代码页的问题。  
  删除数据库：drop database dbname  
2、表：  
  创建新表：  
   create table tabname(col1 type1 [not null] [primary key],col2 type2 [not null],..)  
  根据已有的表创建新表：  
   A：create table tab\_new like tab\_old  
   B：create table tab\_new as select col1,col2… from tab\_old definition only  
修改表：  
增加一个列:  
Alter table tabname add column col type  
注：列增加后将不能删除。DB2中列加上后数据类型也不能改变，唯一能改变的是增加varchar类型的长度。添加主键：  
Alter table tabname add primary key(col)  
删除主键：  
Alter table tabname drop primary key(col)  
 删除表：drop table tabname   
3、表空间：  
  创建表空间：create tablespace tbsname pagesize 4k managed by database using (file ‘file’ size)  
  表空间加入容器：alter tablespace tablespace\_name add(file 'filename' size)  
  注：该操作是不可逆的，加入容器后将不能将其删除，因此在加入的时候注意。  
  删除表空间：drop tablespace tbsname  
4、索引：  
  创建索引：create [unique] index idxname on tabname(col….)   
  删除索引：drop index idxname  
  注：索引是不可更改的，想更改必须删除重新建。  
5、视图：  
  创建视图：create view viewname as select statement  
  删除视图：drop view viewname  
    注：视图唯一能修改的是引用类型列，改变列的范围。其他定义好了都不能修改。当视图基于的基表drop后，视图变为无效。  
  
DML—数据库操纵语言，它不隐式地提交当前事务 ，是否提交视环境设定而定。  
  
 SELECT：从表中查询符合数据  
  注：条件中连接的问题，避免出现笛卡儿乘积  
   
 DELETE：删除已有表的数据  
 UPDATE：更新已有表的数据  
 INSERT：向已有表中插入数据  
 注：DELETE，UPDATE和INSERT是否直接提交取决与执行语句所在的环境。  
 在执行时注意事务日志满的情况。  
1、select时，注意索引谓词和非索引谓词，尽量在有索引的列上使用索引谓词。  
谓词类型 可索引注释  
Col∝con Y ∝代表>;,>;=,=,<=,<,但是<>;不是可索引的。  
Col between con1 and con2 Y 在匹配系列中必须是最后的。  
Col in list Y 仅对一个匹配列  
Col is null Y  
Col like ‘xyz%’ Y 模糊匹配%在后面。  
Col like ‘%xyz’ N 模糊匹配%在前面。  
Col1∝Col2 N Col1和col2来自同一个表  
Col∝Expression N 例如：c1（c1+1）/2  
Pred1 and Pred2 Y Pred1和Pred2都是可索引的，指相同索引的列  
Pred1 or Pred2 N 除了（c1=a or c1=b）外，他可以被认为是c1 in（a，b）  
Not Pred1 N 或者任何的等价形式：Not between,Not in,Not like等等。  
使用索引的例子介绍：  
<1>;、单个表上索引查询的介绍 ：   
A：select \*  from t1 where c1 = 10；  
  在c1列上无索引，如何检索  
  在c1列上有索引，如何检索  
B：select \*  from t1 where c1 =10 and c2 between 5 and 10 and c3 like ‘A%’  
 仅在c1列上有索引  
  在c1,c2和c3列上有单独的索引  
  在c1，c2和c3列上有联合索引  
此处可以填加索引匹配规则和高级规则，比较难理解！！  
<2>;、两个或多个表上索引查询的介绍：  
(1)、使用循环嵌套法执行查询  
A：select t1.c1,t1.c2,t2.c3,t2.c4 from t1,t2 where t1.c1=10 and t1.c2 =t2.c3  
介绍查询的方法，看看哪列上最需要索引。  
B：select t1.c1,t1.c2,t2.c3,t2.c4 from t1,t2 where t1.c1=10 and t2.c4 = 10 and t1.c2 =t2.c3  
(2)、使用归并连接执行连接查询  
例：select t1.c1,t1.c2,t2.c3,t2.c4 from t1,t2 where t1.c1=10 and t2.c4 = 10 and t1.c2 =t2.c3  
(3)、三个以上的表进行连接查询采取的规则或者方法  
（2）、GROUP BY：实现简单分组的功能，当用group by时，查询选择列中除了在group by中出现的和常量外，其他的列上要用分组函数。可以使用一些分组函数实现一些列不在group by中出现，min，max等。  
（3）、HAVING的使用，对一些分组列进行条件判断。  
（4）、ORDER BY子句使得SQL在显示查询结果时将各返回行按顺序排列，返回行的排列顺序由ORDER BY 子句指定的表达式的值确定。  
2、DELETE：从表中删除记录  
语法格式：  
DELETE FROM tablename WHERE (conditions)   
3、INSERT：向表中插入记录  
语法格式：  
INSERT INTO tablename (col1,col2,…) VALUES (value1,value2,…);  
INSERT INTO tablename (col1,col2,…) VALUES (value1, value2,…), (value1, value2,…),……  
Insert不会等待任何程序，不会导致锁定。  
4、UPDATE：  
语法格式：  
UPDATE tabname SET (col1=values1,col2=values2,…) WHERE (conditions);  
注：update的速度比较慢，要在相应列上建立索引。  
  
DCL—数据控制语言  
  
 GRANT—授予用户权限  
 REVOKE—撤消用户权限  
 COMMIT—提交事务，可以使数据库的修改永久化   
 ROLLBACK—回滚事务，消除上一个COMMIT命令后的所做的全部修改，使得数据库的内容恢复到上一个COMMIT执行后的状态.。  
  
1、GRANT：所有着或者管理员把访问权限赋给其他用户  
语法格式：  
grant [all privileges|privileges,….] on tabname | viewname  to [public|user,….]   
2、REVOKE:取消某一用户的某中访问权限  
语法格式：  
Revoke [all privileges|privileges,….] on tabname | viewname  from [public|user,….]   
注：不能取消实例级别的用户的任何权限，他们不是通过grant授权的，是通过组实现的权限。  
3、COMMIT:把事务中所做的修改永久化记录到数据库。  
语法格式：  
 commit [work]  
4、ROLLBACK:将上次提交以来所做的更改全部撤消。  
语法格式：  
Rollback [work]  
  
高级sql简单介绍  
一、查询间使用运算词  
A： UNION 运算符  
UNION 运算符通过组合其他两个结果表（例如 TABLE1 和 TABLE2）并消去表中任何重复行而派生出一个结果表。当 ALL 随 UNION 一起使用时（即 UNION ALL），不消除重复行。两种情况下，派生表的每一行不是来自 TABLE1 就是来自 TABLE2。   
B： EXCEPT 运算符  
EXCEPT 运算符通过包括所有在 TABLE1 中但不在 TABLE2 中的行并消除所有重复行而派生出一个结果表。当 ALL 随 EXCEPT 一起使用时 (EXCEPT ALL)，不消除重复行。   
C： INTERSECT 运算符  
INTERSECT 运算符通过只包括 TABLE1 和 TABLE2 中都有的行并消除所有重复行而派生出一个结果表。当 ALL 随 INTERSECT 一起使用时 (INTERSECT ALL)，不消除重复行。  
注：使用运算词的几个查询结果行必须是一致的。  
二、外连接  
A、left outer join：  
左外连接（左连接）：结果集几包括连接表的匹配行，也包括左连接表的所有行。  
B：right outer join:  
右外连接(右连接)：结果集既包括连接表的匹配连接行，也包括右连接表的所有行。  
C：full outer join：  
全外连接：不仅包括符号连接表的匹配行，还包括两个连接表中的所有记录。  
  
注：复合外连接按照从左到右的顺序执行连接，左边连接的结果集和右边连接  
三、超级分组和移动函数  
A：grouping sets：用来在单个sql中形成多级分组。  
例：select company\_id,node\_id,count(customer\_id) from customer group by  grouping sets(company\_id,node\_id)  
B：rollup：可以在单个数据库操作中形成多个分组。  
 例：select company\_id,node\_id,count(customer\_id) from customer group by  rollup(company\_id,node\_id)  
注：rollup操作不是可交换的操作，指定用户组的顺序是很重要的。  
C：cube： 生成分组表中分组的所有组合。  
 例：select company\_id,node\_id,count(customer\_id) from customer group by  cube(company\_id,node\_id)  
D：over：移动函数可以帮助实现移动的数据分析  
  Select date,avg(qty) over(order by date rows between 1 preceding and 1 following) as values from sale  
  
附录：常用函数介绍  
1、类型转化函数：  
  转化为数字类型的：  
decimal, double, Integer, smallint,real  
Hex(arg):转化为参数的16进制表示。  
  转化为字符串类型的：  
char, varchar  
Digits(arg):返回arg的字符串表示法，arg必须为decimal。  
  转化为日期时间的：  
date, time,timestamp  
2、时间日期：  
year, quarter, month, week, day, hour, minute ,second  
dayofyear(arg):返回arg在年内的天值  
Dayofweek(arg):返回arg在周内的天值  
days(arg):返回日期的整数表示法，从0001-01-01来的天数。   
midnight\_seconds(arg):午夜和arg之间的秒数。  
Monthname(arg):返回arg的月份名。  
Dayname(arg):返回arg的星期。  
3、字符串函数：  
length,lcase, ucase, ltrim, rtrim  
Coalesce(arg1,arg2….):返回参数集中第一个非null参数。  
Concat (arg1,arg2):连接两个字符串arg1和arg2。  
insert(arg1,pos,size,arg2):返回一个，将arg1从pos处删除size个字符，将arg2插入该位置。  
left(arg,length):返回arg最左边的length个字符串。  
locate(arg1,arg2,<pos>;):在arg2中查找arg1第一次出现的位置，指定pos，则从arg2的pos处开始找arg1第一次出现的位置。  
posstr(arg1,arg2):返回arg2第一次在arg1中出现的位置。  
repeat(arg1 ,num\_times):返回arg1被重复num\_times次的字符串。  
replace(arg1,arg2,arg3):将在arg1中的所有arg2替换成arg3。  
right(arg,length):返回一个有arg左边length个字节组成的字符串。  
space(arg):返回一个包含arg个空格的字符串。  
substr(arg1,pos,<length>;):返回arg1中pos位置开始的length个字符，如果没指定length，则返回剩余的字符。  
4、数学函数：   
Abs, count, max, min, sum  
Ceil(arg):返回大于或等于arg的最小整数。  
Floor(arg):返回小于或等于参数的最小整数。  
Mod(arg1,arg2):返回arg1除以arg2的余数，符号与arg1相同。  
Rand():返回1到1之间的随机数。  
Power(arg1,arg2):返回arg1的arg2次方。  
Round(arg1,arg2):四舍五入截断处理，arg2是位数，如果arg2为负，则对小数点前的数做四舍五入处理。  
Sigh(arg):返回arg的符号指示符。-1,0,1表示。  
truncate(arg1,arg2):截断arg1，arg2是位数，如果arg2是负数，则保留arg1小数点前的arg2位。  
  
5、其他：  
nullif(arg1,arg2)：如果2个参数相等，则返回null，否则，返回参数1

1.数据库查询：Select用法  
Select [Top(数值)] 字段列表 From 数据表 [Where 条件] [Order by 字段] [asc或desc]

2.添加数据：Insert Into用法  
Insert Into 数据表(字段1,字段2,字段3,…) Values(字段1的值,字段2的值,字段3的值,…)

3.删除数据：Delete用法  
Delete From 数据表 [Where 条件]

4.更新数据：Update用法  
Update 数据表 Set 字段1=字段值1, 字段2=字段值2, … [Where 条件]

5.建数据表：Create Table用法  
Create Table 数据表名(字段1名称,字段1类型,字段2名称 字段2类型, …)

Access数据表常用数据类型:Text,Char(Number),Memo,Number,Int,Date/Time,Logical,OLEObject

自动编号字段添加例子:create table aaa(id int identity (1, 1) not null,abc varchar(25) null)

6.改数据表：Alter Table用法  
添加字段：Alter Table 数据表名 Add Column 字段名 字段类型

删除字段：Alter Table 数据表名 Drop Column 字段名

7.删数据表：Drop Table用法

Drop Table 数据表名