创建emp表

create table emp (

empno int,

ename string,

job string,

mgr int,

hiredate string,

sal double,

comm double,

deptno int

)

row format delimited fields terminated by ‘\t’;

创建dept表

create table dept(

deptno int,

dname string,

loc int

)

row format delimited fields terminated by ‘\t’;

将数据文件移动到对应的表所在的HDFS路径下，完成数据映射

开始对数据进行操作：

1基本查询

select \* from emp;

select deptno, dname from dept;

select count(\*) cnt from emp;

select sum(sal) sum\_sal from emp;

select \* from emp limit 5;

2where条件查询

select \* from emp where sal = 5000;

select \* from emp where sal between 500 and 1000;

select \* from emp where comm is null;

select \* from emp where sal in (1500, 5000);

3like和rlike

select \* from emp where sal like ‘2%’;

select \* from emp where sal like ‘\_2%’;

select \* from emp where sal like ‘[2]’;

4Group by语句

select t.deptno, avg(t.sal) avg\_sal from emp t group by t.deptno;

select t.deptno, t.job, max(t.sal) max\_sal from emp t group by t.deptno, t.job;

5having语句

select deptno, avg(sal) from emp group by deptno;

select deptno, avg(sal) avg\_sal from emp group by deptno having avg\_sal > 2000;

6order by 语句

select \* from emp order by sal desc;

select ename, deptno, sal from emp order by deptno, sal;

7sort by 语句

set mapreduce.job.reduces=3;

set mapreduce.job.reduces;

select \* from emp sort by empno desc;

将查询结果导入到文件中

insert overwrite local directory ‘/root/sortby-result’ select \* from emp sort by deptno desc;

8Distribute by

set mapreduce.job.reduces=3;

insert overwrite local directory ‘/root/distribute-result’

select \* from distribute by deptno sort by empno desc;

9Cluster by

以下两种写法等价

select \* from emp cluster by deptno;

select \* from emp distribute by deptno sort by deptno;

10 join

当前版本中，Hive只支持等值连接

select e.empno, e.ename, d.deptno, d.dname

from emp e join dept d on e.deptno=d.deptno ;

select e.empno, e.ename, d.deptno

from emp e left join dept d on e.deptno=d.deptno ;

select e.empno, e.ename, d.deptno

from emp e right join dept d on e.deptno=d.deptno ;

满外连接：

select e.empno, e.ename, d.deptno

from emp e full join dept d on e.deptno=d.deptno ;

select e.empno, e.ename, d.deptno

from emp e full join dept d on e.deptno=d.deptno

where d.deptno=20;