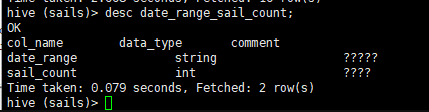
实验 4-4 Hive ql 实现价格、 时间段销量分析

建议课时： 60 分钟

一、 实验目的  
λ 掌握 Hql 语句的编写；  
λ 了解时间段和价格两个指标的分析方法；  
二、 实验环境  
Dsight 智慧实验室的 hadoop 环境  
三、 实验步骤  
本节实验所做内容如下红色标注：  
本节实验主要是通过 Hive ql 语句实现各时间段、 各价格区间手机销量的统  
计。  
具体实验步骤如下：  
1. 划分时间段统计手机销量  
（1） 新建时间段手机销量表  
表结构设计：  
drop table if exists date\_range\_sail\_count;  
create table date\_range\_sail\_count  
( date\_range string comment '交易时间段',  
sail\_count int comment '手机销量'  
) comment '各时间段手机销量表';



（2） 编辑 Hql 语句关联查询数据并插入新表关联手机销售信息表和用户行为表各时间段销量表字段、 数据：

insert into date\_range\_sail\_count

select deal\_time,count(\*) as count

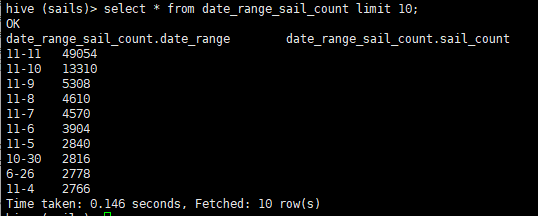
from goods\_sail\_tb g

join user\_action\_tb u on u.goods\_id=g.goods\_id

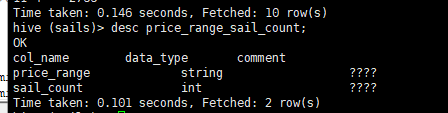
group by deal\_time

order by count desc

limit 10;



2. 划分价格区间统计手机销量  
（1） 新建价格区间中间表， 存储手机销量、 价格区间  
<1000 1000 元以下  
[1000,2000] 1000-2000 元  
[2001,3000] 2000-3000 元  
[3001-5000] 3000-5000 元  
[5001-8000] 5000-8000 元  
[8001-10000] 8000-10000 元  
>10000 10000 元以上  
（2） 在中间表基础上计算销量并存储到价格区间手机销量结果表  
表结构设计：  
drop table if exists price\_range\_sail\_count;  
create table price\_range\_sail\_count  
(  
price\_range string comment '价格区间',  
sail\_count int comment '手机销量'  
) comment '各价格区间手机销量表';



建立中间表：

create table price\_range\_sail\_count\_mid as

select

case

when goods\_price<1000 then '1000 元以下'

when goods\_price>=1000 and goods\_price<2000 then '1000-2000元'

when goods\_price>=2000 and goods\_price<3000 then '2000-3000元'

when goods\_price>=3000 and goods\_price<4000 then '3000-5000元'

when goods\_price>=5000 and goods\_price<8000 then '5000-8000元'

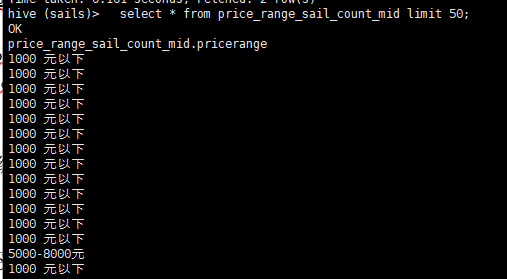
when goods\_price>=8000 and goods\_price<1000 then '8000-10000元'

when goods\_price>=10000 then '10000元以上 '

else '1000 元以下'

end as pricerange

from goods\_sail\_tb



（3） 编辑 Hql 语句计算并查询数据插入结果表

insert into price\_range\_sail\_count

select pricerange,count(\*)

from price\_range\_sail\_count\_mid

group by pricerange;

