实验 4-1 单品销售分析

建议课时： 60 分钟

一、 实验目的

掌握 hive 统计分析的 Hql 语句的编写；

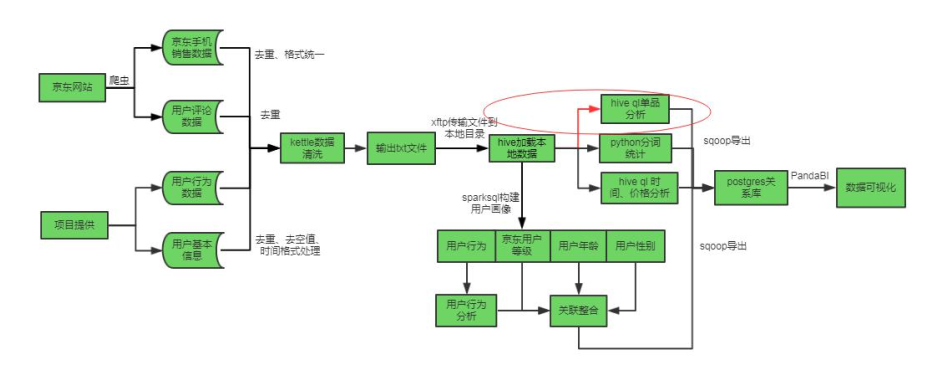
了解单品分析的分析指标；

掌握 hive 多表关联分析方法；  
二、 实验环境

Dsight 智慧实验室的 hadoop 环境

三、 实验步骤

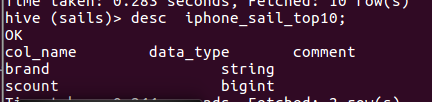
本节实验所做内容如下红色标注：

  
本节实验主要是通过 Hive ql 语句进行手机

具体实验步骤如下：  
1. 进入 hadoop 环境， 启动 hive,进入自定义数据库  
2. 基于手机销售新表， 统计各手机品牌销量 top10：  
Hql 语句统计 TopN 并建表， 表字段包含：手机品牌、 销售量：

建表语句：

create table iphone\_sail\_top10(brand string,scount bigint) ;



插入数据：

insert into iphone\_sail\_top10

SELECT brand,count(\*) as count FROM

goods\_sail\_tb

GROUP BY

brand

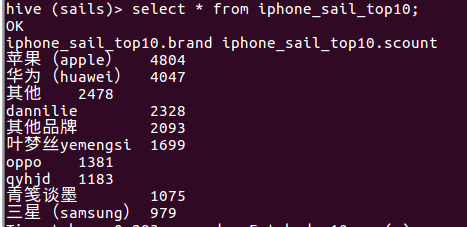
ORDER BY count DESC

limit 1,10



查询

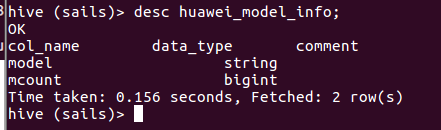
select \* from iphone\_sail\_top10



3. 统计某品牌的各型号手机的销量  
（1） 统计华为手机品牌下各单品销量 Top20：（方便后续大屏展示）  
选择手机销售信息表中的手机品牌（查询条件为华为）、 销售量字段  
作为新表字段， 查询并建表。

建表

create table huawei\_model\_info(model string,mcount bigint) ;



插入语句：

Insert into Huawei\_model\_info

SELECT brand,count(\*) as count FROM goods\_sail\_tb where brand like ‘华为’

group BY

brand

ORDER BY count DESC

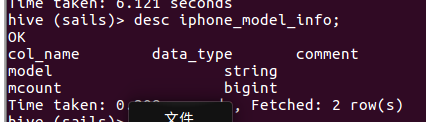
limit 20



（2） 统计苹果手机品牌下各单品销量 Top20（方便后续大屏展示）：选择手机销售信息表中的手机品牌（查询条件为苹果和 Apple）、 销售量字段作为新表字段， 查询并建表。

建表语句：

create table iphone\_model\_info(model string,mcount bigint) ;



插入数据：

insert into iphone\_model\_info

SELECT brand,count(\*) as count FROM goods\_sail\_tb

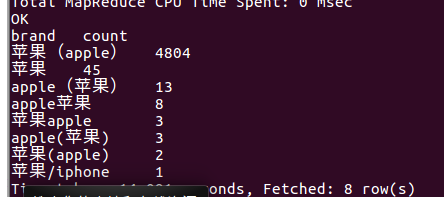
where brand like '%苹果%'

group BY

brand

ORDER BY count DESC

limit 20



4. 统计用户各年龄段手机销量  
（1） 关联查询新用户信息表、 用户评论表、 手机销量表生成新 hive 表，统计各年龄段的手机销量。 生成的新表包含的字段：用户年龄段、 单品手机类型、 销量数量

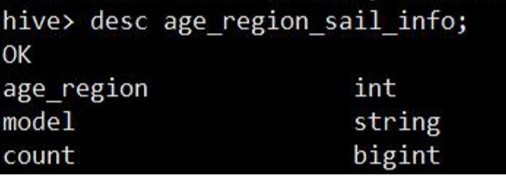
建表语句：

create table age\_region\_sail\_info(

age\_region int,

model string,

count bigint)



插入数据：

Insert into table age\_region\_sail\_info

select i.age\_region\_alias,g.brand ,count(\*) as count

from user\_comment\_tb as c

join user\_info\_new\_tb as i on i.user\_id=c.user\_id

join goods\_sail\_tb as g on g.goods\_id =c.goods\_id

group by i.age\_region\_alias,g.brand

5. 统计各地区用户手机销量  
（1） 关联用户信息表、 用户评价表、 手机销量表生成新表， 新表包含的字段：用户地区、 手机单品类型、 销售量

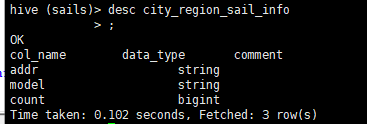
建表：

create table city\_region\_sail\_info(

addr string,

model string,

count bigint)



插入数据：

insert into city\_region\_sail\_info

select substr(i.addr,1,3) ,g.brand ,count(\*) as count

from user\_comment\_tb as c

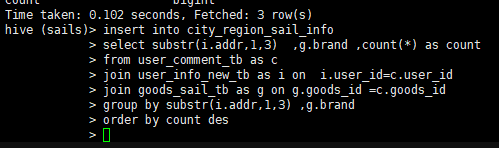
join user\_info\_new\_tb as i on i.user\_id=c.user\_id

join goods\_sail\_tb as g on g.goods\_id =c.goods\_id

group by substr(i.addr,1,3) ,g.brand

order by count desc

limit 10



查询：

select \* from city\_region\_sail\_info limit 10;

