

描述

题目：现在运营举办了一场比赛，收到了一些参赛申请，表数据记录形式如下所示，现在运营想要统计每个性别的用户分别有多少参赛者，请取出相应结果

示例：表user\_submit

|  |  |  |
| --- | --- | --- |
| device\_id | profile | blog\_url |
| 2138 | 180cm,75kg,27,male | http:/url/bigboy777 |
| 3214 | 165cm,45kg,26,female | http:/url/kittycc |
| 6543 | 178cm,65kg,25,male | http:/url/tiger |
| 4321 | 171cm,55kg,23,female | http:/url/uhksd |
| 2131 | 168cm,45kg,22,female | http:/urlsydney |

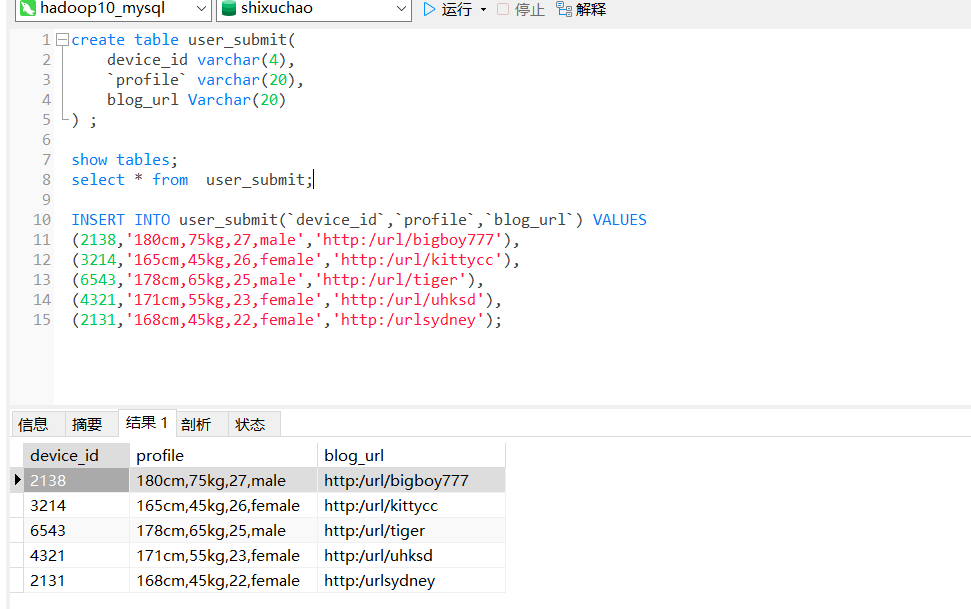
device\_id ：用户id

profile ：用户信息

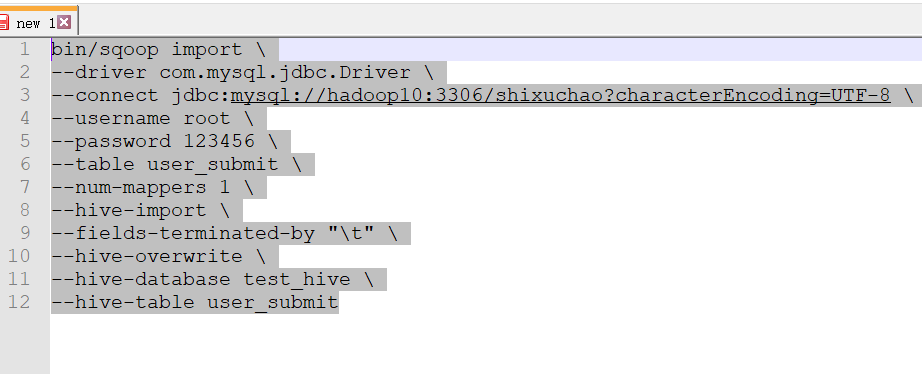
blog\_url ：博客地址

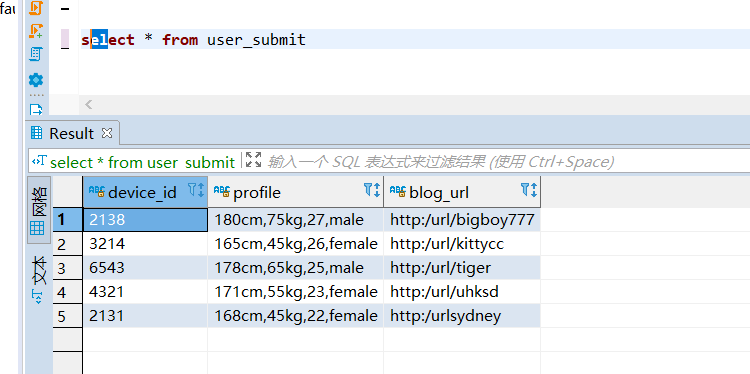
需求：

1. 在MySQL数据库中的创建 表user\_submit ，并将上述数据插入表中。

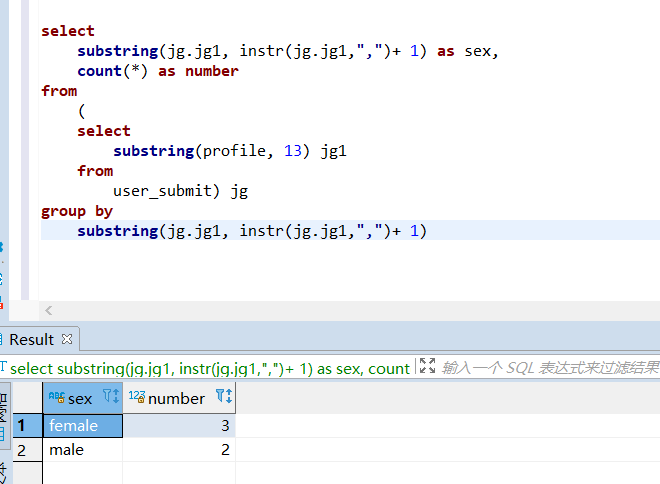


1. 通过sqoop将数据传输至hive





1. 在hive中进行SQL编写与执行
   1. 计算每个性别参加比赛的人数



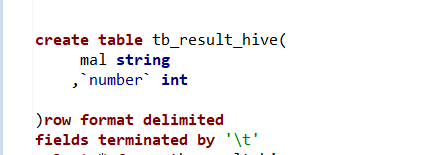
查询应返回以下结果：

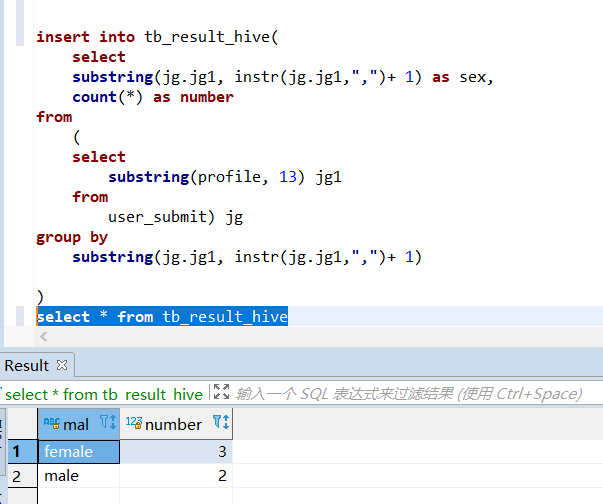
gender number

male 2

female 3

* 1. 将结果放入新的结果表tb\_result\_hive中 ，要求表中的字段分隔符为 “\t”





1. 通过sqoop将hive计算出的结果表tb\_result\_hive传输至MySQL（新建表tb\_result）中。

