

强化训练

面试题

```
year  month  amount
1991,1,1.1
1991,2,1.2
1991,3,1.3
1991,4,1.4
1992,1,2.1
1992,2,2.2
1992,3,2.3
1992,4,2.4
```

讲表格转化为下面结果

```
year m1  m2  m3  m4
1991 1.1 1.2 1.3 1.4
1992 2.1 2.2 2.3 2.4
```

```
SELECT
    year,
    MAX(CASE WHEN month = 1 THEN amount END) AS m1,
    MAX(CASE WHEN month = 2 THEN amount END) AS m2,
    MAX(CASE WHEN month = 3 THEN amount END) AS m3,
    MAX(CASE WHEN month = 4 THEN amount END) AS m4
FROM
    tes1
GROUP BY
    year;
```

```
CREATE TABLE tes1 (
    year int,
    month int,
    amount FLOAT
)row format delimited fields terminated by ",";

load data local inpath '/home/data/year.txt' into table tes1;
```

源代码

```

SELECT
    year,
    MAX(CASE WHEN month = 1 THEN amount END) AS m1,
    MAX(CASE WHEN month = 2 THEN amount END) AS m2,
    MAX(CASE WHEN month = 3 THEN amount END) AS m3,
    MAX(CASE WHEN month = 4 THEN amount END) AS m4
FROM
    tes1
GROUP BY
    year;

```

```

cpx - WindTerm
会话 编辑 搜索 选择 转到 查看 模式 工具 窗口 帮助
1.cpy 2.cpy:2
ssh cpy
2079 Time taken: 3.391 seconds, Fetched: 2 row(s)
2080 hive (mydb)> SELECT
2081 > year,
2082 > MAX(CASE WHEN month = 1 THEN amount END) AS m1,
2083 > MAX(CASE WHEN month = 2 THEN amount END) AS m2,
2084 > MAX(CASE WHEN month = 3 THEN amount END) AS m3,
2085 > MAX(CASE WHEN month = 4 THEN amount END) AS m4
2086 > FROM
2087 > tes1
2088 > GROUP BY
2089 > year;
2090 Automatically selecting local only mode for query
2091 WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e. spark, tez) or using Hive 1.X releases.
2092 Query ID = root_20230627195148_2eb80d40-3352-465e-bf31-cf503bed47ab
2093 Total jobs = 1
2094 Launching Job 1 out of 1
2095 Number of reduce tasks not specified. Estimated from input data size: 1
2096 In order to change the average load for a reducer (in bytes):
2097   set hive.exec.reducers.bytes.per.reducer=<number>
2098 In order to limit the maximum number of reducers:
2099   set hive.exec.reducers.max=<number>
2100 In order to set a constant number of reducers:
2101   set mapreduce.job.reduces=<number>
2102 Job running in-process (local Hadoop)
2103 2023-06-27 19:51:50,375 Stage-1 map = 100%, reduce = 100%
2104 Ended Job = job_local40369273_0025
2105 MapReduce Jobs Launched:
2106 Stage-Stage-1: HDFS Read: 175982 HDFS Write: 163719 SUCCESS
2107 Total MapReduce CPU Time Spent: 0 msec
2108 OK
2109 year    m1    m2    m3    m4
2110 1991    1.1    1.2    1.3    1.4
2111 1992    2.1    2.2    2.3    2.4

```

正在讲话: 华迪信息;

19:51 100.0% 1.51/3.68 GiB

就绪 远程模式 窗口 0x0 行 2113 字符 13 linux 2023/6/27 11:51 WindTerm Issues 锁屏

日志处理

时间	接口	ip地址
2016-11-09 11: 22: 05	/api/user/login	110.23.5.33
2016-11-09 11: 23: 10	/api/user/detail	57.3.2.16
.....		
2016-11-09 23: 59: 40	/api/user/login	200.6.5.166

数据集:

```

2016-11-09 14:22:05 /api/user/login 110.23.5.33 1
2016-11-09 11:23:10 /api/user/detail 57.3.2.16
2016-11-09 14:59:40 /api/user/login 200.6.5.166 1
2016-11-09 14:22:05 /api/user/login 110.23.5.34 1
2016-11-09 14:22:05 /api/user/login 110.23.5.34 1
2016-11-09 14:22:05 /api/user/login 110.23.5.34 1
2016-11-09 11:23:10 /api/user/detail 57.3.2.16
2016-11-09 23:59:40 /api/user/login 200.6.5.166
2016-11-09 14:22:05 /api/user/login 110.23.5.34 1
2016-11-09 11:23:10 /api/user/detail 57.3.2.16
2016-11-09 23:59:40 /api/user/login 200.6.5.166
2016-11-09 14:22:05 /api/user/login 110.23.5.35 1
2016-11-09 14:23:10 /api/user/detail 57.3.2.16
2016-11-09 23:59:40 /api/user/login 200.6.5.166
2016-11-09 14:59:40 /api/user/login 200.6.5.166 1
2016-11-09 14:59:40 /api/user/login 200.6.5.166 1

```

```

create table ip(
time string,
interface string,
ip string)
row format delimited fields terminated by '\t';
load data local inpath '/home/data/ip.dat' into table ip;

select ip, count(ip) cip from ip
where interface = "/api/user/login" and time between '2016-11-09 14:00:00' and
'2016-11-09 15:00:00'
group by ip
order by cip DESC
limit 3;

```

求11月9号下午14点（14-15点），访问api/user/login接口的，top3的ip地址以及次数

源代码

```

select ip, count(ip) cip from ip
where interface = "/api/user/login" and time between '2016-11-09 14:00:00' and
'2016-11-09 15:00:00'
group by ip
order by cip DESC
limit 3;

```

结果

```

Time taken: 3.523 seconds, Fetched: 3 row(s)
hive (mydb)> select ip, count(ip) cip from ip
> where interface = "/api/user/login" and time between '2016-11-09 14:00:00' and '2016-11-09 15:00:00'
> group by ip
> order by cip DESC
> limit 3;
Automatically selecting local only mode for query

```

```
cpy - WindTerm
正在讲话:
会话 编辑 搜索 选择 转到 查看 模式 工具 窗口 帮助
1.cpy 2.cpy:2
ssh cpy
841 Number of reduce tasks not specified. Estimated from input data size: 1
842 In order to change the average load for a reducer (in bytes):
843   set hive.exec.reducers.bytes.per.reducer=<number>
844 In order to limit the maximum number of reducers:
845   set hive.exec.reducers.max=<number>
846 In order to set a constant number of reducers:
847   set mapreduce.job.reduces=<number>
848 Job running in-process (local Hadoop)
849 2023-06-27 19:08:31,602 Stage-1 map = 100%, reduce = 100%
850 Ended Job = job_local1812666940_0009
851 Launching Job 2 out of 2
852 Number of reduce tasks determined at compile time: 1
853 In order to change the average load for a reducer (in bytes):
854   set hive.exec.reducers.bytes.per.reducer=<number>
855 In order to limit the maximum number of reducers:
856   set hive.exec.reducers.max=<number>
857 In order to set a constant number of reducers:
858   set mapreduce.job.reduces=<number>
859 Selecting local mode for task: Stage-2
860 Job running in-process (local Hadoop)
861 2023-06-27 19:08:33,629 Stage-2 map = 100%, reduce = 100%
862 Ended Job = job_local1943789254_0010
863 MapReduce Jobs Launched:
864 Stage-Stage-1:  HDFS Read: 42310 HDFS Write: 26480 SUCCESS
865 Stage-Stage-2:  HDFS Read: 50137 HDFS Write: 37165 SUCCESS
866 Total MapReduce CPU Time Spent: 0 msec
867 OK
868 ip      cip
869 110.23.5.34      4
870 200.6.5.166      3
871 110.23.5.35      1
872 Time taken: 3.41 seconds, Fetched: 3 row(s)
873 hive (mydb)>
```

三连冠

找出全部夺得3连冠的队伍

team,year

活塞,1990
公牛,1991
公牛,1992
公牛,1993
火箭,1994
火箭,1995
公牛,1996
公牛,1997
公牛,1998
马刺,1999
湖人,2000
湖人,2001
湖人,2002
马刺,2003
活塞,2004
马刺,2005
热火,2006
马刺,2007
凯尔特人,2008

湖人,2009
湖人,2010
小牛,2011
热火,2012
热火,2013
马刺,2014
勇士,2015
骑士,2016
勇士,2017
勇士,2018
猛龙,2019
湖人,2020
雄鹿,2021
勇士,2022
掘金,2023

```
create table ball(  
team string,  
year int  
)row format delimited fields terminated by ',';  
  
load data local inpath "/home/data/qiapi.dat" into table ball;
```

查询代码:

提示: rank值的除了排名外, 还有什么含义

源代码

```
SELECT team, count(*) AS TripleCount FROM(  
SELECT team FROM(  
SELECT team, year, ROW_NUMBER()  
OVER (PARTITION BY team ORDER BY year) AS row_num  
FROM ball  
) subquery  
GROUP BY team, year - row_num  
HAVING COUNT(*) >= 3  
) subquery2  
GROUP by team;
```

结果

```
Time taken: 1.079 seconds, Fetched: 2 row(s)
hive (mydb)> SELECT team, count(*) AS TripleCount FROM(
> SELECT team FROM(
> SELECT team, year, ROW_NUMBER()
> OVER (PARTITION BY team ORDER BY year) AS row_num
> FROM ball
> ) subquery
> GROUP BY team, year - row_num
> HAVING COUNT(*) >= 3
> ) subquery2
> GROUP by team;
```

```
cpy - WindTerm
会话 编辑 搜索 选择 转到 查看 模式 工具 窗口 帮助
1.cpy 2.cpy:2
ssh cpy
2127 Launching Job 1 out of 2
2128 Number of reduce tasks not specified. Estimated from input data size: 1
2129 In order to change the average load for a reducer (in bytes):
2130   set hive.exec.reducers.bytes.per.reducer=<number>
2131 In order to limit the maximum number of reducers:
2132   set hive.exec.reducers.max=<number>
2133 In order to set a constant number of reducers:
2134   set mapreduce.job.reduces=<number>
2135 Job running in-process (local Hadoop)
2136 2023-06-27 19:56:35,606 Stage-1 map = 100%, reduce = 100%
2137 Ended Job = job_local603654031_0026
2138 Launching Job 2 out of 2
2139 Number of reduce tasks not specified. Estimated from input data size: 1
2140 In order to change the average load for a reducer (in bytes):
2141   set hive.exec.reducers.bytes.per.reducer=<number>
2142 In order to limit the maximum number of reducers:
2143   set hive.exec.reducers.max=<number>
2144 In order to set a constant number of reducers:
2145   set mapreduce.job.reduces=<number>
2146 Selecting local mode for task: Stage-2
2147 Job running in-process (local Hadoop)
2148 2023-06-27 19:56:37,613 Stage-2 map = 100%, reduce = 100%
2149 Ended Job = job_local811047928_0027
2150 MapReduce Jobs Launched:
2151 Stage-Stage-1:  HDFS Read: 177116 HDFS Write: 164646 SUCCESS
2152 Stage-Stage-2:  HDFS Read: 187483 HDFS Write: 178429 SUCCESS
2153 Total MapReduce CPU Time Spent: 0 msec
2154 OK
2155 team    triplecount
2156 公牛    2
2157 湖人    1
2158 Time taken: 3.417 seconds, Fetched: 2 row(s)
2159 hive (mydb)>
```

正在讲话: [Microphone icon]

19:56 0.0% M 1.52/3.68 GiB

就绪 远程模式 窗口 0x0 行 2159 字符 13 linux 2023/6/27 11:56 WindTerm Issues 锁屏

