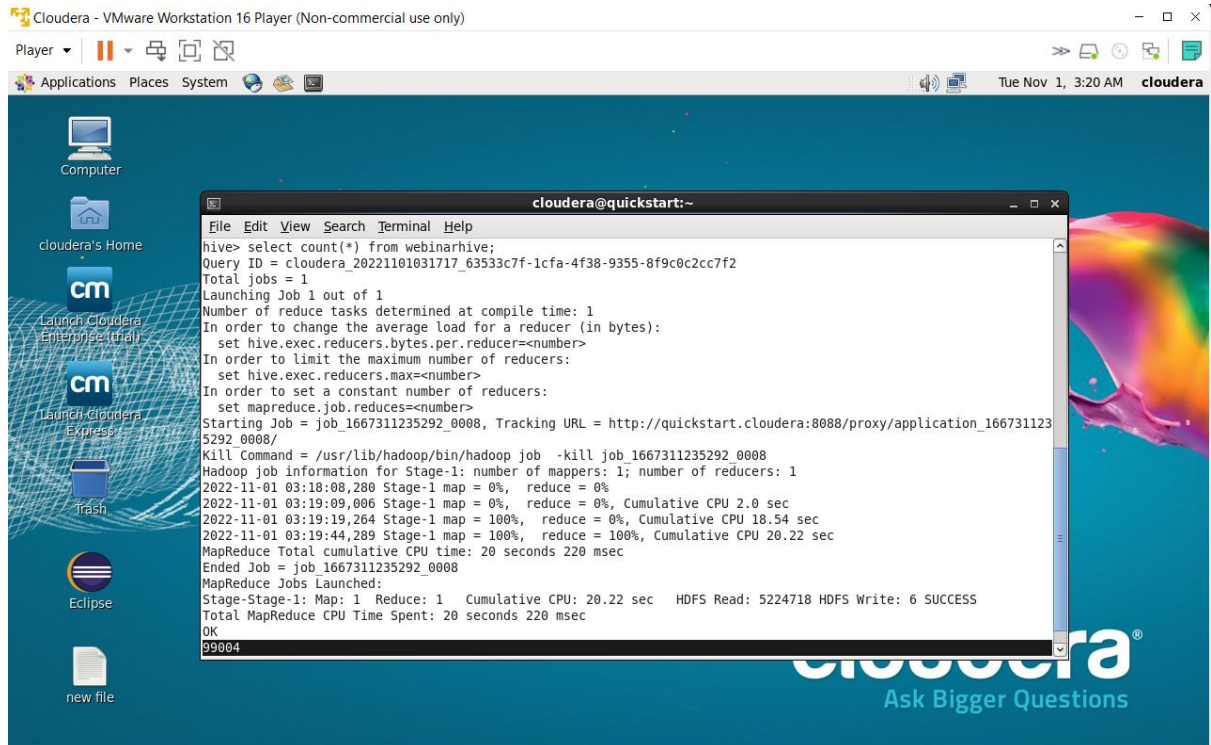


1. Find out total number of users in dataset.

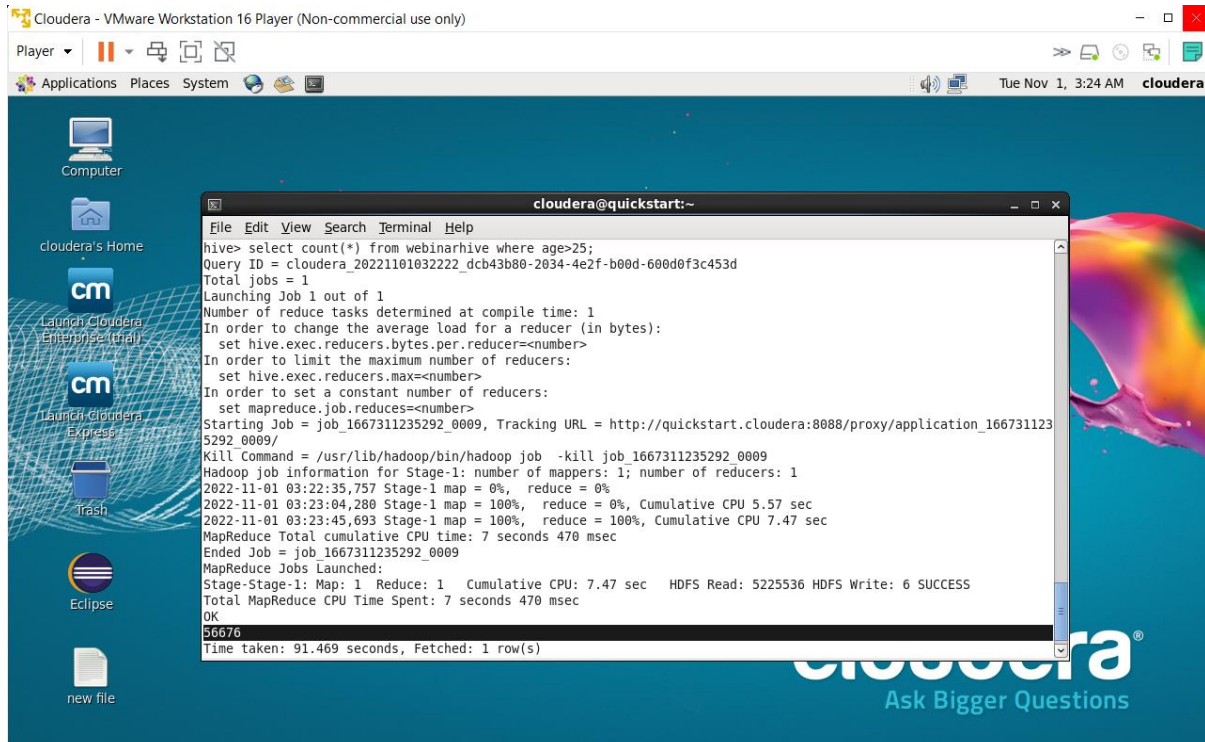


The screenshot shows a VMware Workstation 16 Player window titled "Cloudera - VMware Workstation 16 Player (Non-commercial use only)". Inside the VM, a terminal window titled "cloudera@quickstart:~" displays the execution of a Hive query. The query is "select count(\*) from webinarhive;". The output shows the job is launched and completed successfully. The final result is 99004.

```
File Edit View Search Terminal Help
hive> select count(*) from webinarhive;
Query ID = cloudera_20221101031717_63533c7f-1cfa-4f38-9355-8f9c0c2c7f2
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks determined at compile time: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1667311235292_0008, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1667311235292_0008/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1667311235292_0008
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-11-01 03:18:08,280 Stage-1 map = 0%, reduce = 0%
2022-11-01 03:19:09,006 Stage-1 map = 0%, reduce = 0%, Cumulative CPU 2.0 sec
2022-11-01 03:19:19,264 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 18.54 sec
2022-11-01 03:19:44,289 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 20.22 sec
MapReduce Total cumulative CPU time: 20 seconds 220 msec
Ended Job = job_1667311235292_0008
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 20.22 sec HDFS Read: 5224718 HDFS Write: 6 SUCCESS
Total MapReduce CPU Time Spent: 20 seconds 220 msec
OK
99004
```

Analysis Result:- 99004 users are in dataset.

2. Find out the number of facebook user above the age 25.

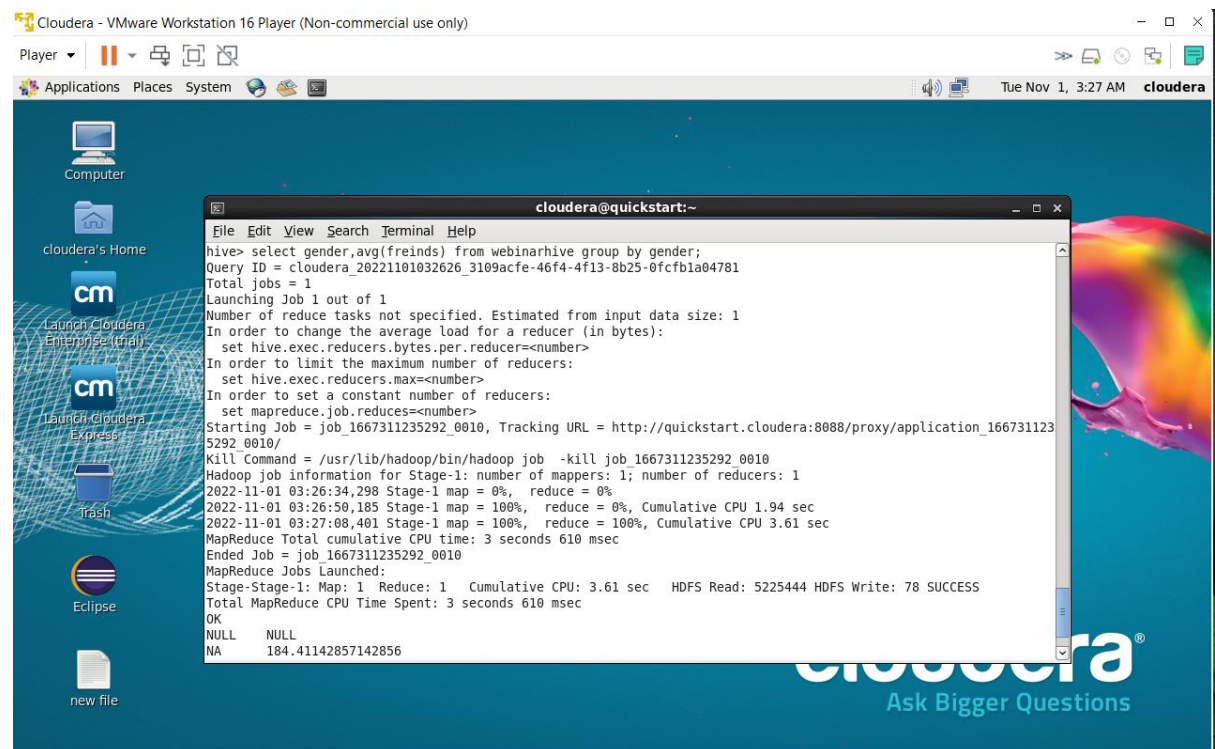


The screenshot shows a VMware Workstation 16 Player window titled "Cloudera - VMware Workstation 16 Player (Non-commercial use only)". Inside the VM, a terminal window titled "cloudera@quickstart:~" displays the execution of a Hive query. The query is "select count(\*) from webinarhive where age>25;". The output shows the job is launched and completed successfully. The final result is 56676.

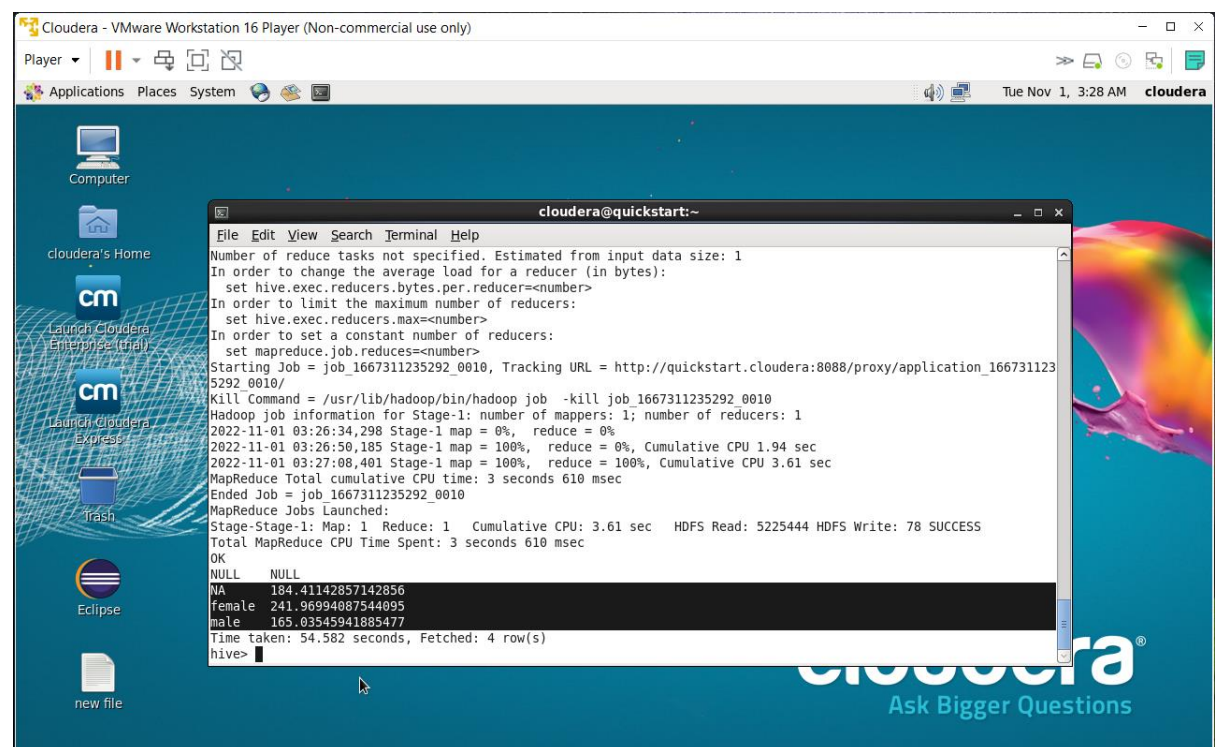
```
File Edit View Search Terminal Help
hive> select count(*) from webinarhive where age>25;
Query ID = cloudera_20221101032222_dcb43b80-2034-4e2f-b00d-600d0f3c453d
Total jobs = 1
Launching Job 1 out of 1
Number of reduce tasks determined at compile time: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1667311235292_0009, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1667311235292_0009/
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1667311235292_0009
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-11-01 03:22:35,757 Stage-1 map = 0%, reduce = 0%
2022-11-01 03:23:04,280 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 5.57 sec
2022-11-01 03:23:45,693 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 7.47 sec
MapReduce Total cumulative CPU time: 7 seconds 470 msec
Ended Job = job_1667311235292_0009
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 7.47 sec HDFS Read: 5225536 HDFS Write: 6 SUCCESS
Total MapReduce CPU Time Spent: 7 seconds 470 msec
OK
56676
Time taken: 91.469 seconds, Fetched: 1 row(s)
```

Analysis Result :- 56676

### 3. DO male facebook user tend to have more friends, or female users?



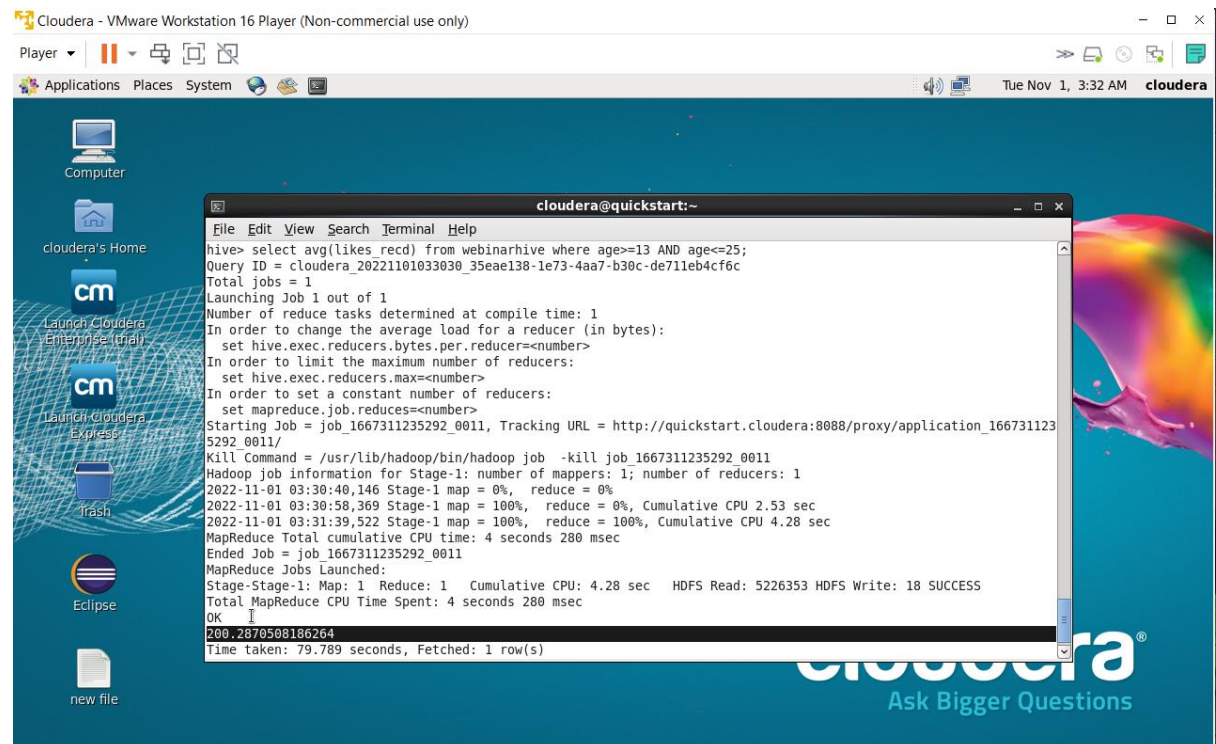
```
cloudera@quickstart:~  
File Edit View Search Terminal Help  
hive> select gender, avg(freinds) from webinarhive group by gender;  
Query ID = cloudera_20221101032626_3109acfe-46f4-4f13-8b25-0fcfb1a04781  
Total jobs = 1  
Launching Job 1 out of 1  
Number of reduce tasks not specified. Estimated from input data size: 1  
In order to change the average load for a reducer (in bytes):  
  set hive.exec.reducers.bytes.per.reducer=<number>  
In order to limit the maximum number of reducers:  
  set hive.exec.reducers.max=<number>  
In order to set a constant number of reducers:  
  set mapreduce.job.reduces=<number>  
Starting Job = job_1667311235292_0010, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1667311235292_0010/  
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1667311235292_0010  
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1  
2022-11-01 03:26:34,298 Stage-1 map = 0%, reduce = 0%  
2022-11-01 03:26:50,185 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.94 sec  
2022-11-01 03:27:08,401 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 3.61 sec  
MapReduce Total cumulative CPU time: 3 seconds 610 msec  
Ended Job = job_1667311235292_0010  
MapReduce Jobs Launched:  
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 3.61 sec HDFS Read: 5225444 HDFS Write: 78 SUCCESS  
Total MapReduce CPU Time Spent: 3 seconds 610 msec  
OK  
NULL NULL  
NA 184.41142857142856
```



```
cloudera@quickstart:~  
File Edit View Search Terminal Help  
Number of reduce tasks not specified. Estimated from input data size: 1  
In order to change the average load for a reducer (in bytes):  
  set hive.exec.reducers.bytes.per.reducer=<number>  
In order to limit the maximum number of reducers:  
  set hive.exec.reducers.max=<number>  
In order to set a constant number of reducers:  
  set mapreduce.job.reduces=<number>  
Starting Job = job_1667311235292_0010, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1667311235292_0010/  
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1667311235292_0010  
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1  
2022-11-01 03:26:34,298 Stage-1 map = 0%, reduce = 0%  
2022-11-01 03:26:50,185 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.94 sec  
2022-11-01 03:27:08,401 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 3.61 sec  
MapReduce Total cumulative CPU time: 3 seconds 610 msec  
Ended Job = job_1667311235292_0010  
MapReduce Jobs Launched:  
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 3.61 sec HDFS Read: 5225444 HDFS Write: 78 SUCCESS  
Total MapReduce CPU Time Spent: 3 seconds 610 msec  
OK  
NULL NULL  
NA 184.41142857142856  
female 241.96994087544095  
male 165.03545941885477  
Time taken: 54.582 seconds, Fetched: 4 row(s)  
hive>
```

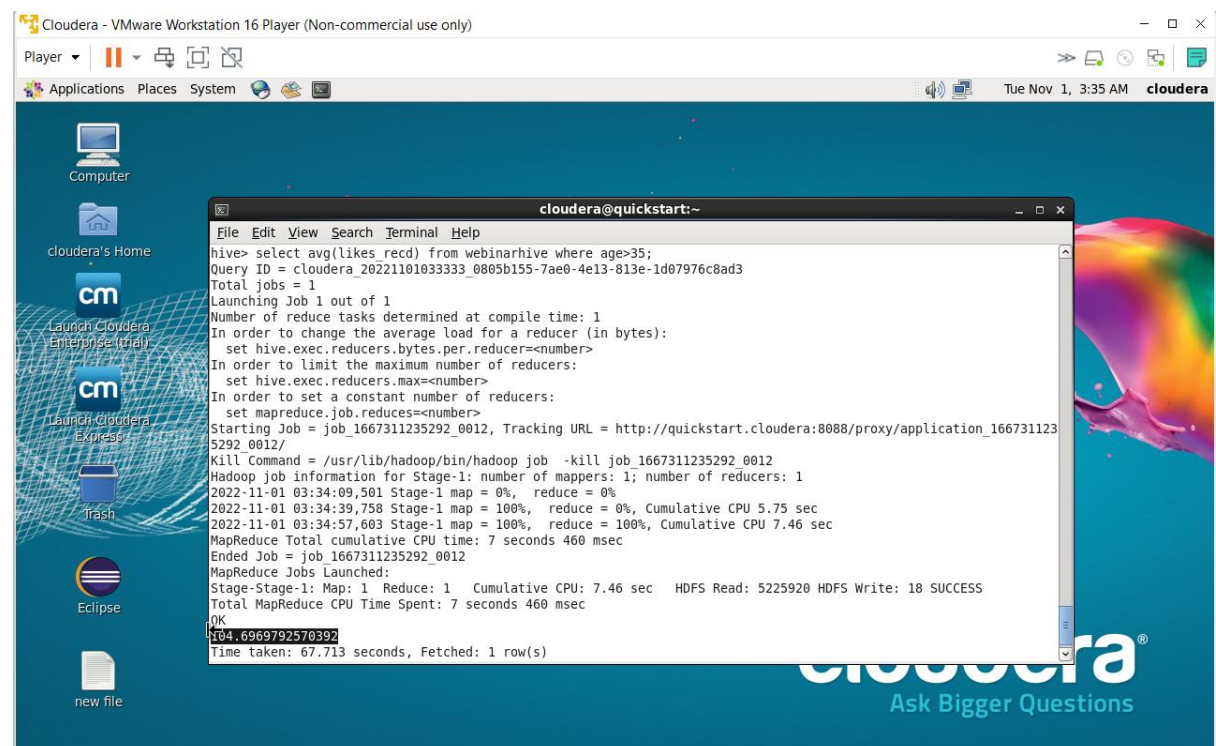
**Analysis Result:- Female have more Friends.**

#### 4. How many likes do young people receive on facebook opposed to older members



The screenshot shows a terminal window in a Cloudera VM. The user has executed a Hive query: `select avg(likes_recd) from webinarhive where age >= 13 AND age <= 25;`. The output shows the job is launched and completed successfully. The final result is `200.2870508186264`.

```
cloudera@quickstart:~  
File Edit View Search Terminal Help  
hive> select avg(likes_recd) from webinarhive where age >= 13 AND age <= 25;  
Query ID = cloudera_20221101033030_35eae138-1e73-4aa7-b30c-de711eb4cf6c  
Total jobs = 1  
Launching Job 1 out of 1  
Number of reduce tasks determined at compile time: 1  
In order to change the average load for a reducer (in bytes):  
  set hive.exec.reducers.bytes.per.reducer=<number>  
In order to limit the maximum number of reducers:  
  set hive.exec.reducers.max=<number>  
In order to set a constant number of reducers:  
  set mapreduce.job.reduces=<number>  
Starting Job = job_1667311235292_0011, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1667311235292_0011/  
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1667311235292_0011  
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1  
2022-11-01 03:30:40,146 Stage-1 map = 0%, reduce = 0%  
2022-11-01 03:30:58,369 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.53 sec  
2022-11-01 03:31:39,522 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 4.28 sec  
MapReduce Total cumulative CPU time: 4 seconds 280 msec  
Ended Job = job_1667311235292_0011  
MapReduce Jobs Launched:  
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 4.28 sec HDFS Read: 5226353 HDFS Write: 18 SUCCESS  
Total MapReduce CPU Time Spent: 4 seconds 280 msec  
OK  
200.2870508186264  
Time taken: 79.789 seconds, Fetched: 1 row(s)
```



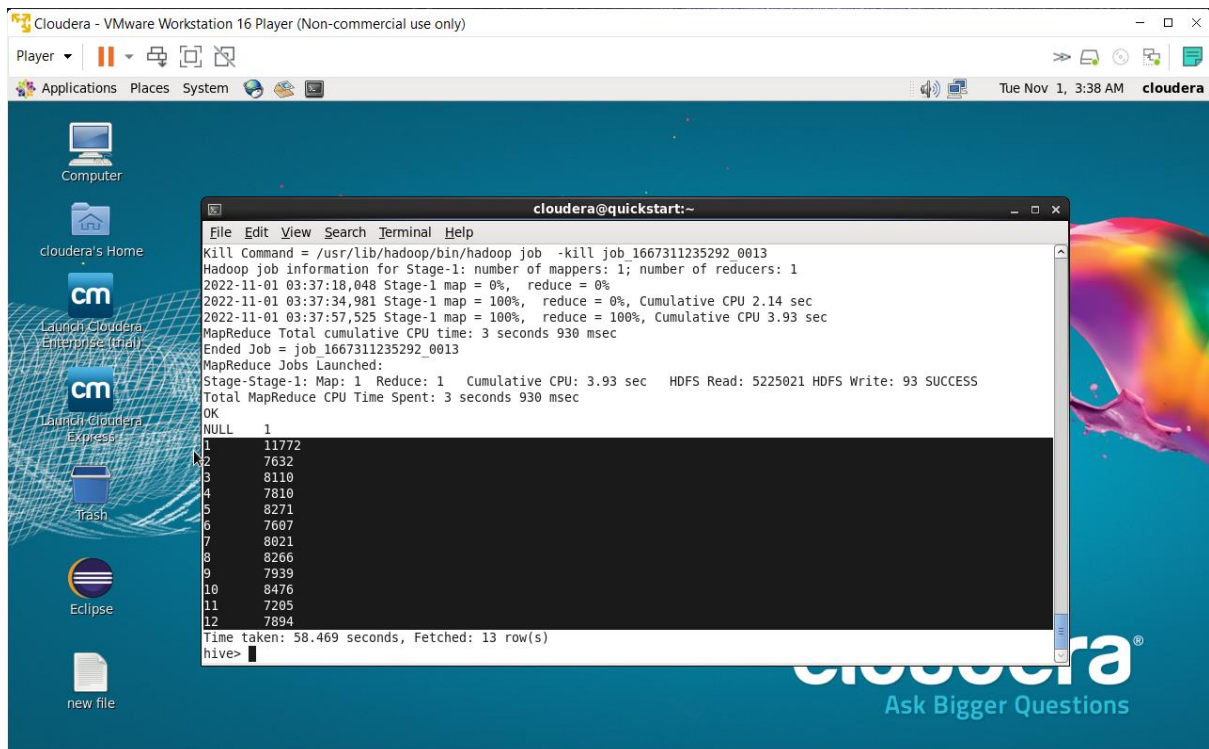
The screenshot shows a terminal window in a Cloudera VM. The user has executed a Hive query: `select avg(likes_recd) from webinarhive where age > 35;`. The output shows the job is launched and completed successfully. The final result is `104.6960792570392`.

```
cloudera@quickstart:~  
File Edit View Search Terminal Help  
hive> select avg(likes_recd) from webinarhive where age > 35;  
Query ID = cloudera_20221101033333_0805b155-7ae0-4e13-813e-1d07976c8ad3  
Total jobs = 1  
Launching Job 1 out of 1  
Number of reduce tasks determined at compile time: 1  
In order to change the average load for a reducer (in bytes):  
  set hive.exec.reducers.bytes.per.reducer=<number>  
In order to limit the maximum number of reducers:  
  set hive.exec.reducers.max=<number>  
In order to set a constant number of reducers:  
  set mapreduce.job.reduces=<number>  
Starting Job = job_1667311235292_0012, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1667311235292_0012/  
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1667311235292_0012  
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1  
2022-11-01 03:34:09,501 Stage-1 map = 0%, reduce = 0%  
2022-11-01 03:34:39,758 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 5.75 sec  
2022-11-01 03:34:57,603 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 7.46 sec  
MapReduce Total cumulative CPU time: 7 seconds 460 msec  
Ended Job = job_1667311235292_0012  
MapReduce Jobs Launched:  
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 7.46 sec HDFS Read: 5225920 HDFS Write: 18 SUCCESS  
Total MapReduce CPU Time Spent: 7 seconds 460 msec  
OK  
104.6960792570392  
Time taken: 67.713 seconds, Fetched: 1 row(s)
```

Analylsis Result : - Young member received more likes than older.



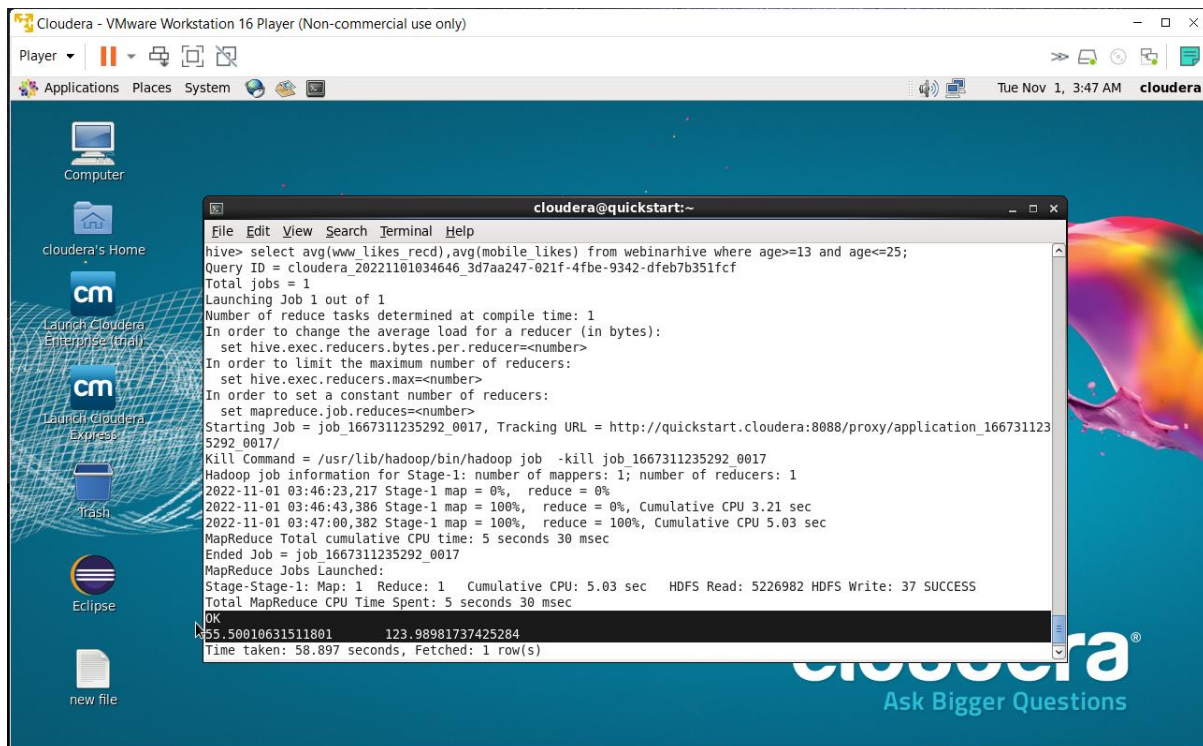
## 5. Find out the count of facebook user for each birthday month



The screenshot shows a VMware Workstation 16 Player window titled "Cloudera - VMware Workstation 16 Player (Non-commercial use only)". The desktop background is the Cloudera logo with the tagline "Ask Bigger Questions". The desktop has several icons: "Computer", "cloudera's Home", "Launch Cloudera Enterprise (trial)", "Launch Cloudera Express", "Trash", "Eclipse", and "new file". A terminal window titled "cloudera@quickstart:~" is open, displaying the following text:

```
File Edit View Search Terminal Help
Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1667311235292_0013
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2022-11-01 03:37:18,048 Stage-1 map = 0%, reduce = 0%
2022-11-01 03:37:34,981 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.14 sec
2022-11-01 03:37:57,525 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 3.93 sec
MapReduce Total cumulative CPU time: 3 seconds 930 msec
Ended Job = job_1667311235292_0013
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 3.93 sec HDFS Read: 5225021 HDFS Write: 93 SUCCESS
Total MapReduce CPU Time Spent: 3 seconds 930 msec
OK
NULL 1
1 11772
2 7632
3 8110
4 7810
5 8271
6 7607
7 8021
8 8266
9 7939
10 8476
11 7205
12 7894
Time taken: 58.469 seconds, Fetched: 13 row(s)
hive>
```

6. Do young members use mobile phones or computer for browsing?



7. Do adult members use mobile phones or computer for browsing?

