**Facebook Data Analysis**

### Create table facebook

To see all csv files available in HDFS :

ls -l

csv file path: hdfs dfs -put /home/mavricbdhnov0433/pseudo\_facebook.csv

[mavricbdhnov0433@ip-10-1-1-204 ~]$ hdfs dfs -put /home/mavricbdhnov0433/pseudo\_facebook.csv

put: `pseudo\_facebook.csv': File exists

[mavricbdhnov0433@ip-10-1-1-204 ~]$

A screenshot of a computer

Description automatically generated

Create database :

Hive> create database khusboosharma;

show databases like '\*khushboo\*';

hive> show databases like '\*khushboo\*';

OK

Time taken: 0.434 seconds

hive> use khusboosharma;

OK

Time taken: 0.04 seconds

Create fb\_data table in khusboosharma database :

Hive> create table fb\_data(id int, age int, day int, year int, month int, gender string, tenure int, friends int,

friend\_init int, likes int, likes\_recd int,mlikes int, mlikes\_recd int, wlikes int, wlikes\_recd int)

row format delimited

fields terminated by ',';

OK

Time taken: 0.435 seconds

**To load the data into Hive:**

LOAD DATA local INPATH 'pseudo\_facebook.csv' OVERWRITE INTO TABLE fb\_data;

**To see table structure:**

describe fb\_data;

**To see the table data:**

select \* from fb\_data limit 50;

**1. Find out the total number of users in this dataset.**

select count(\*) from fb\_data;

A screenshot of a computer

Description automatically generated

Q2: Find out the no of users above the age of 25

select count(\*) from fb\_data where age>25;

A screenshot of a computer

Description automatically generated

Q3: Do male face book users tend to have more friends or female users

select gender,avg(friends) from fb\_data group by gender;

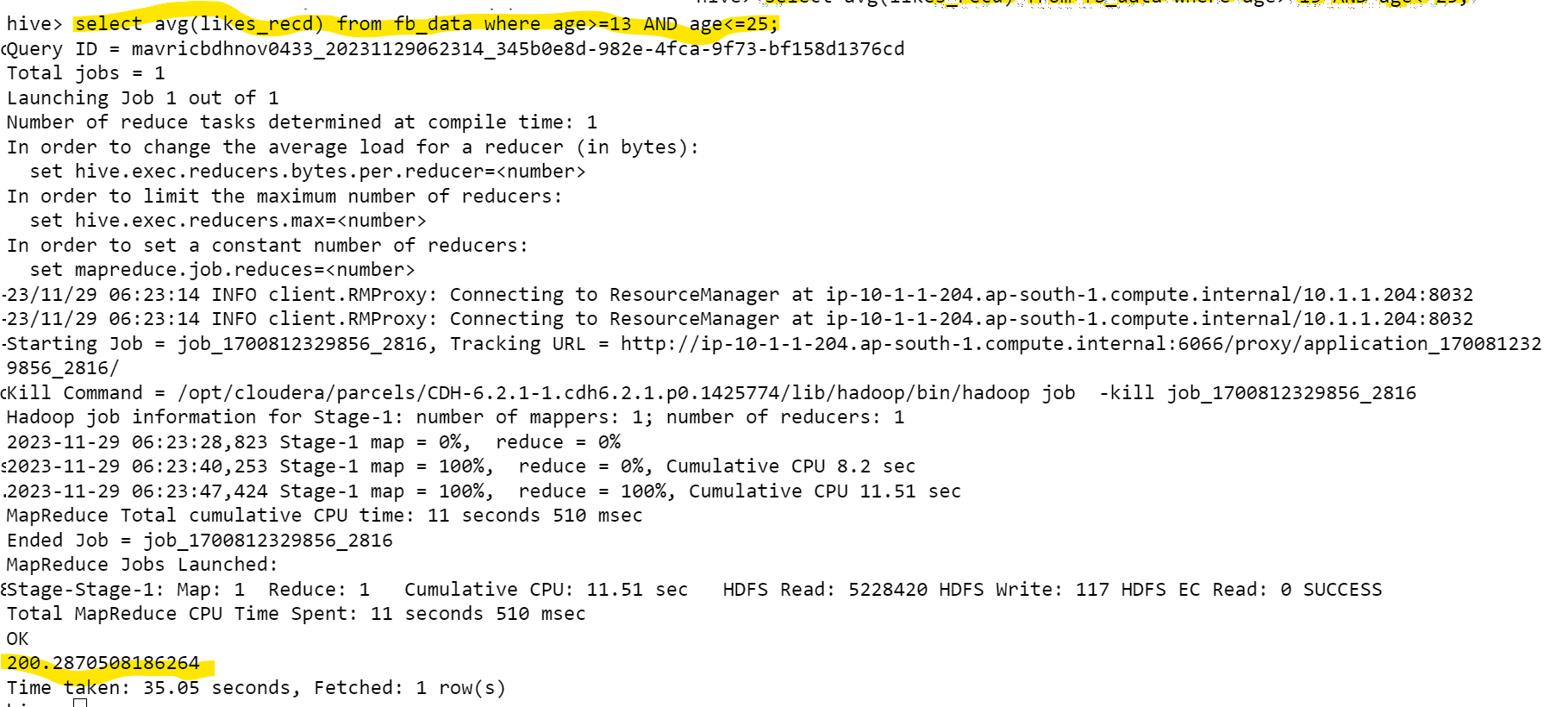
A screenshot of a computer

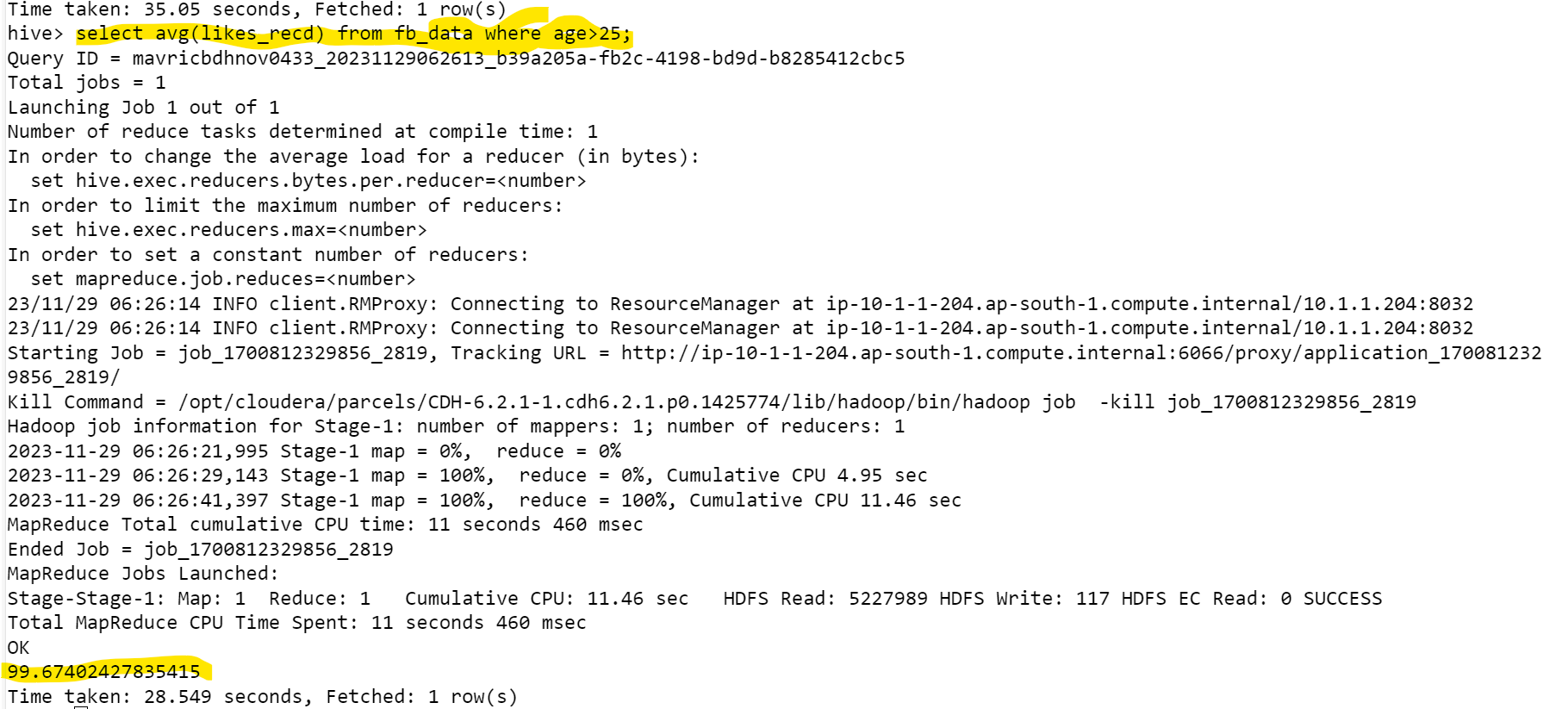
Description automatically generated

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

select avg(likes\_recd) from fb\_data where age>=13 AND age<=25;

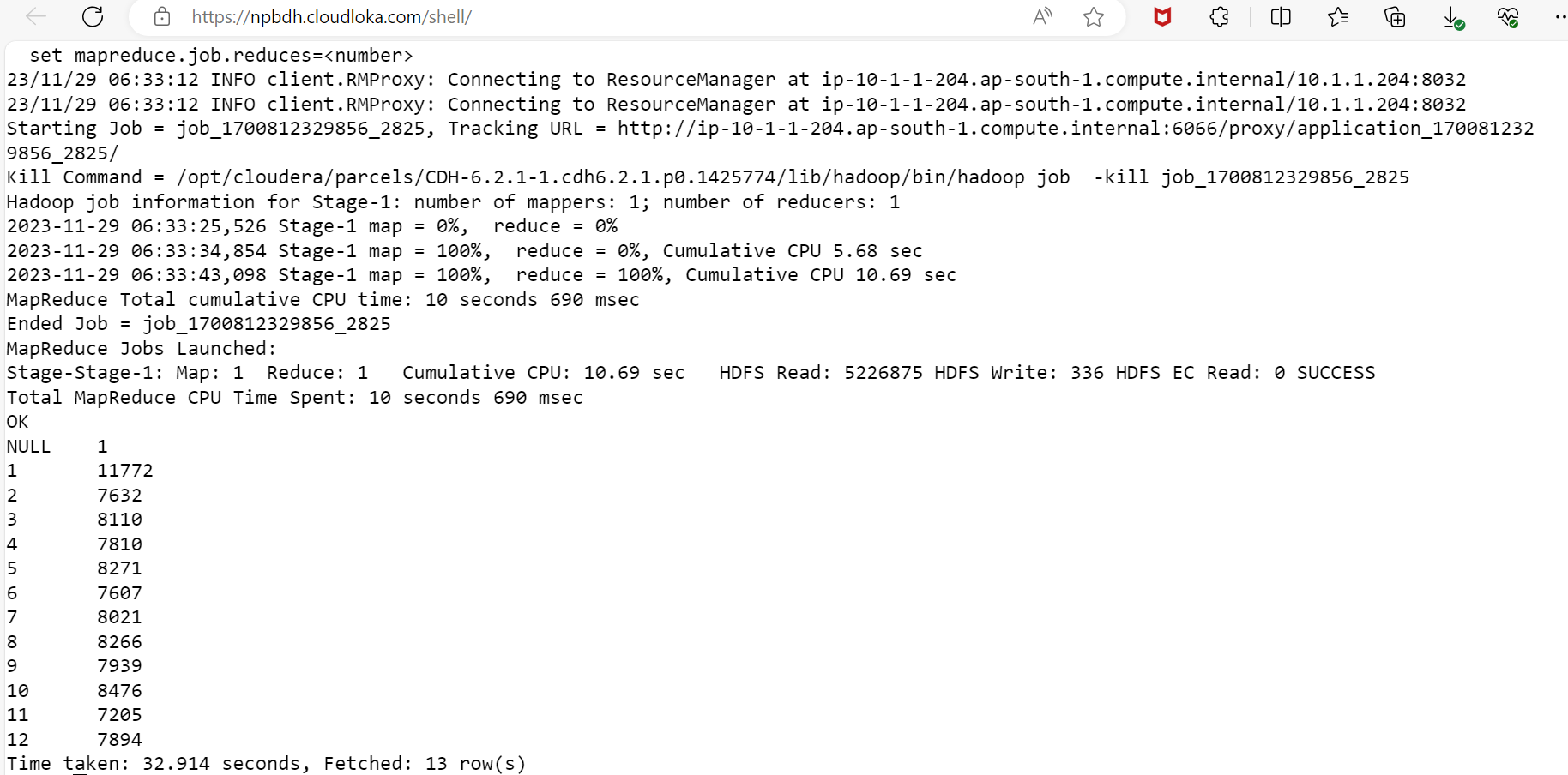
select avg(likes\_recd) from fb\_data where age>25;





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

select month,count(\*) from fb\_data group by month;



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

select avg(mlikes),avg(wlikes) from fb\_data where age>=13 AND age<=25;

A close up of a computer screen

Description automatically generated

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

select avg(mlikes),avg(wlikes) from fb\_data where age>=35;

