***Proof of Concept***

***On***

**Title: FaceBook Data Analysis**

**Submitted for the requirement of**

**Big Data Engineering Course**

BACHELOR OF ENGINEERING

**COMPUTER SCIENCE & ENGINEERING (Big Data and Analytics)**

**CSC-334**

**Semester-6**

****

**Submitted to:**

Ms. Gurpreet Kaur

Project Supervisor

**Submitted by:**

Arman Gupta(18BCS3759)

Vishavdeep Singh Aulakh(18BCS3767)

**ACKNOWLEDGEMENT**

We would like to express our deepest appreciation to all those who provided us the possibility to complete this report. A special gratitude we give to our 5th semester B.D.E project supervisor, Ms. Gurpreet Kaur, whose contribution in stimulating suggestions and encouragement, helped us to coordinate our project and especially in writing this report. Furthermore, we would also like to acknowledge with much appreciation her crucial role, in giving the permission to use all required equipment and the necessary materials to complete the task ‘Analysis of Facebook Dataset’ using Apache Hadoop (Hive), IBM Cognos Insight and Python Programming, and gave suggestion about the task.

**OVERVIEW**

Facebook has not only changed how we communicate to each other, but how we collect data for the benefit of our business. As opposed to big budget ad campaigns that often become ineffective due to no direction, Facebook has refined its advertising mechanism so that target users will see your product and enjoy it. These advancements in online marketing have made it possible to interact when more data is collected from the users, which is opposed to the days of the user data being stored provided little to no avenues of strategy for the marketer. With this being said, here’s a brief look at how Facebook Data Analytics benefit not only how a company invests into marketing… but the effectiveness of their marketing strategies in relations to the customer

**OBJECTIVES**

Brand awareness

Increase overall awareness for your brand by showing ads to people who are more likely to pay attention to them.

Works well with: [ad recall lift](https://en-gb.facebook.com/business/help/1029827880390718?ref=fbb_adobjectives)

Reach

Show ads to the maximum number of people in your audience while staying within your budget. You can also choose to reach only people who are near your business locations.

**COLUMNS AND DATA TYPE :**

**Age int**

**Id 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**

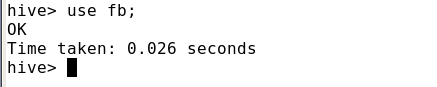
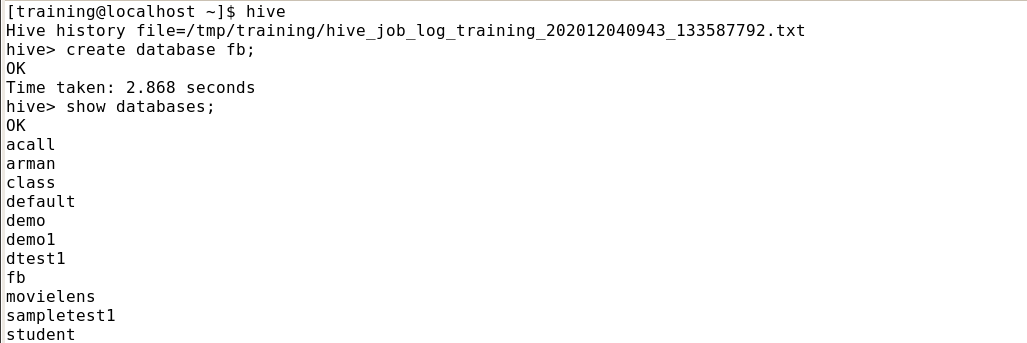
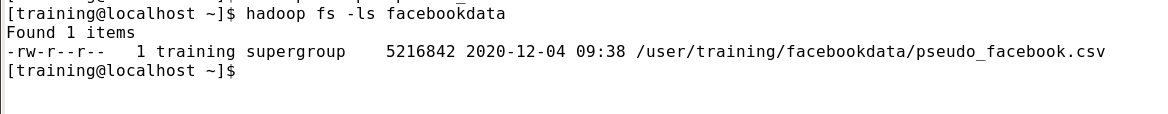
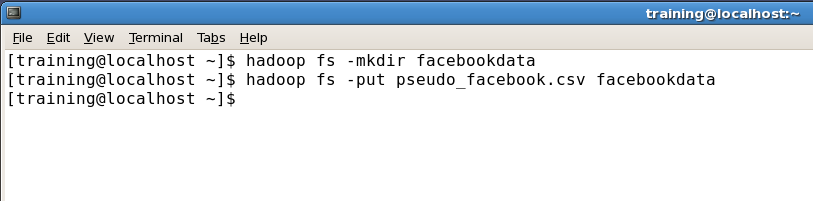
**wLikes int**

**wlikes\_recd int**

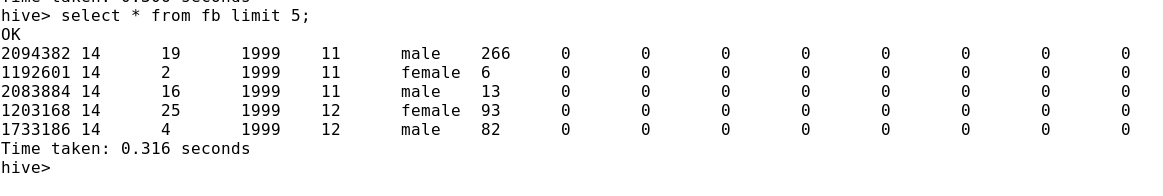
**PROBLEM STATEMENTS:**

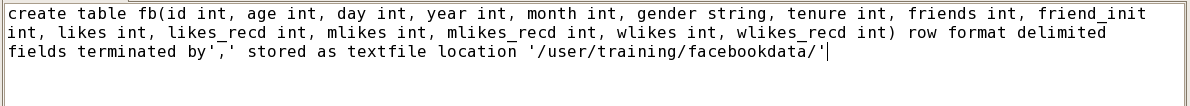
1. Find the total number of users in this dataset.
2. Find out the number of Facebook users above the age of 25.
3. Do male Facebook users tend to have more friends ,or female users?
4. How many likes do young people receive on Facebook opposed to older members
5. Find out the count of Facebook users for each birthday month.
6. Do young members use mobile phones or computers for Facebook browsing?
7. Do adult members use mobile phones or computers for Facebook browsing
8. Visualisation graph for the age wise number of people on Facebook.
9. Visualisation for the number of likes which was received by male and female.
10. Visualisation for the likes received for the age of the users (Male or Female)

**HIVE QUERIES**

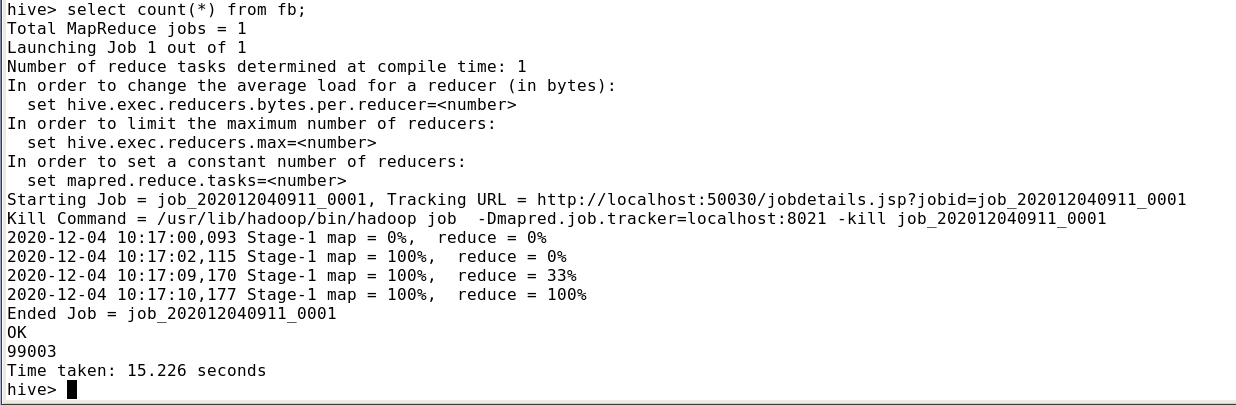
1. Create a directory and copy the data in it. 

2. Creating a database and use it.

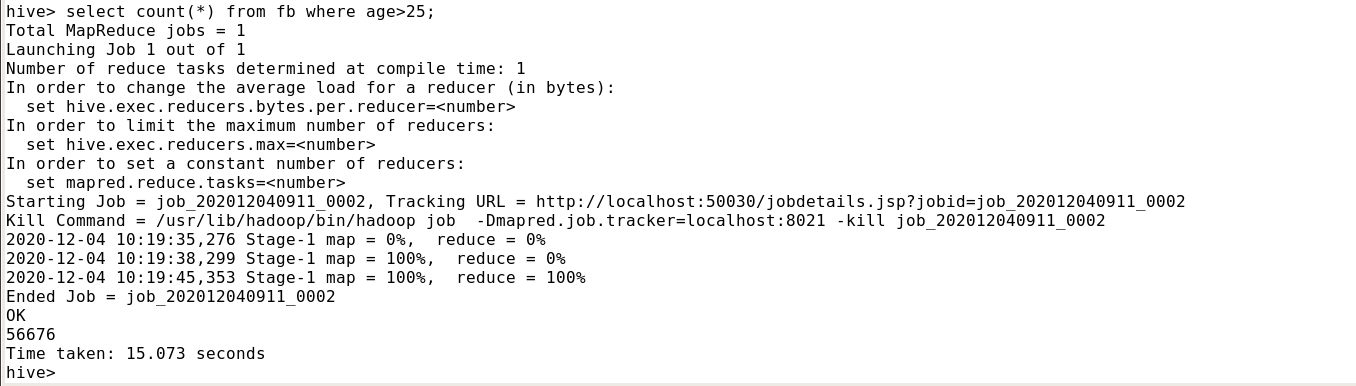
3. Displaying the top 5 rows of uploaded data:

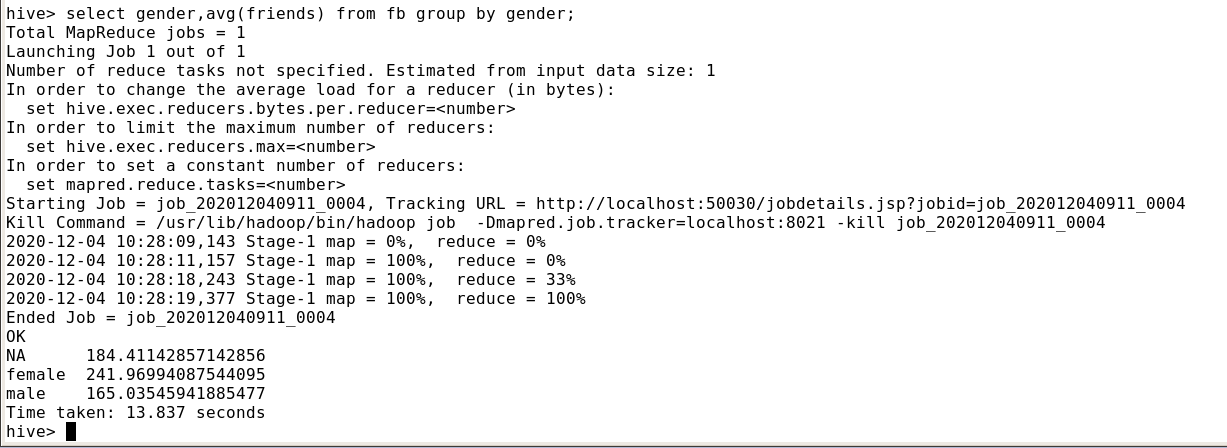
4. Creating a hive table:

**PROBLEM STATEMENT 1:** **Find the total number of users in this dataset.**

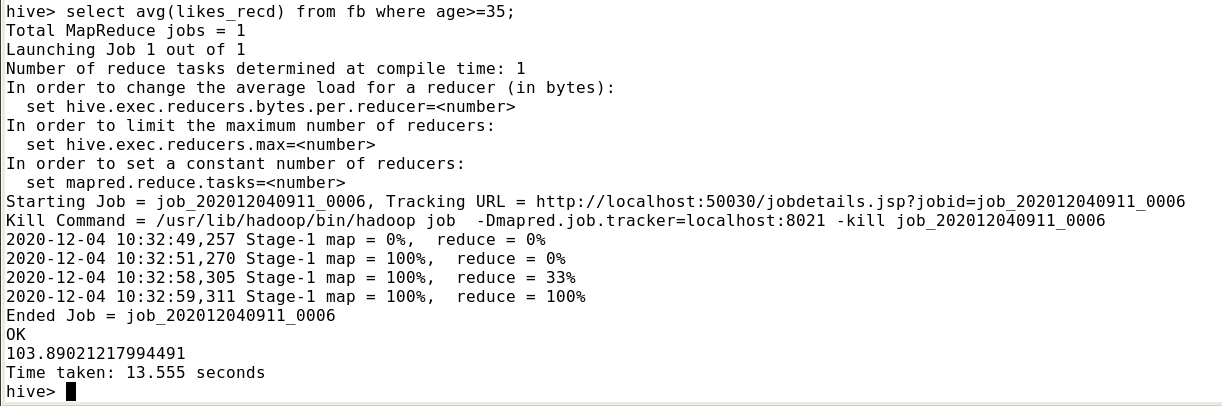
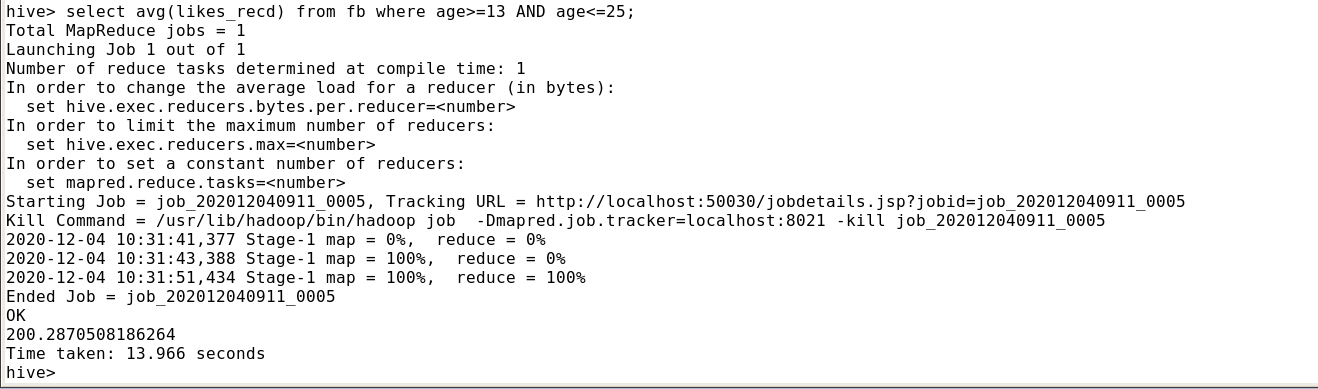
****

**PROBLEM STATEMENT 2: Find out the number of Facebook users above the age of 25.**

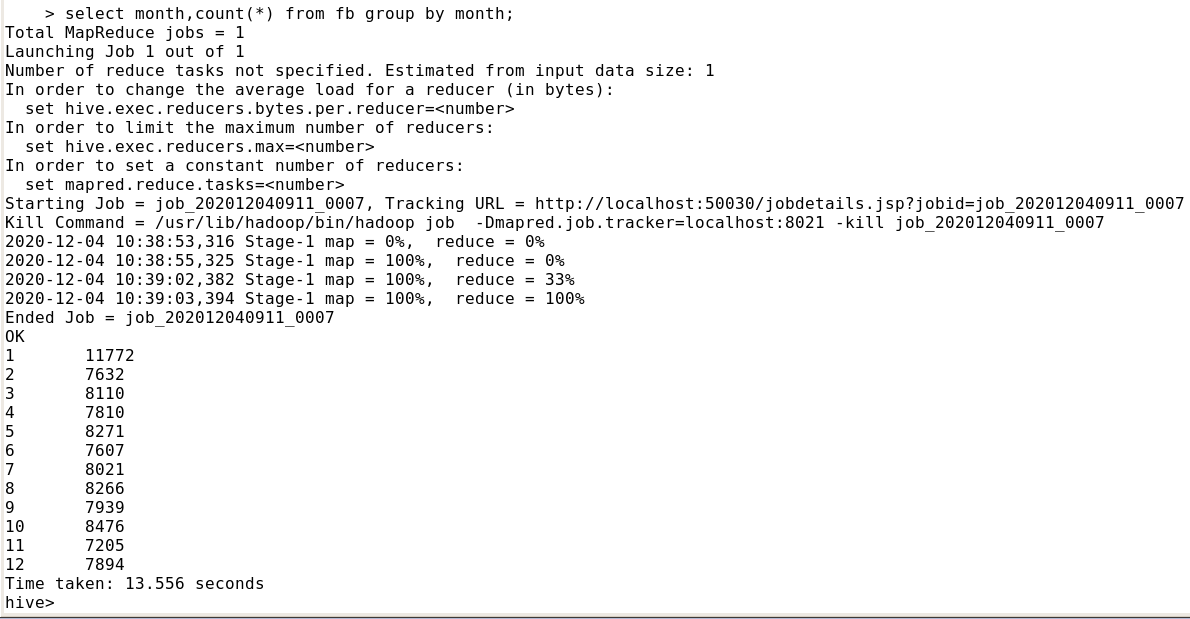
****

**PROBLEM STATEMENT 3:Do male Facebook users tend to have more friends ,or female users?**

This result is as expected and quite obvious female receives more like then male .So brand or product can select the girl or lady who received most of the likes or more socially active on Facebook for brand promotion.

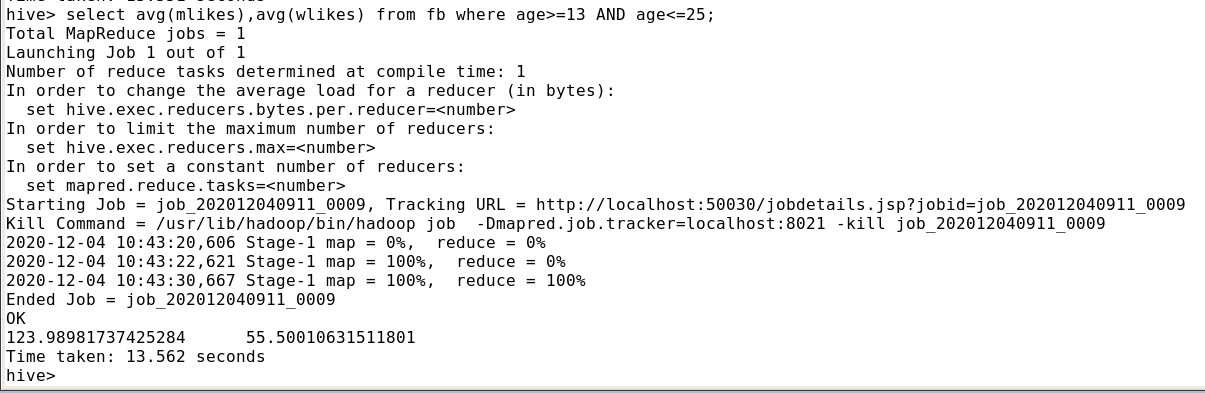
**PROBLEM STATEMENT 4:How many likes do young people receive on Facebook opposed to older members ?**

We’ve use average function as we use in sql. We can also take sum but if any outliers is present in dataset so error can be occur . So the result clearly shows that number of young people is more than the old people .

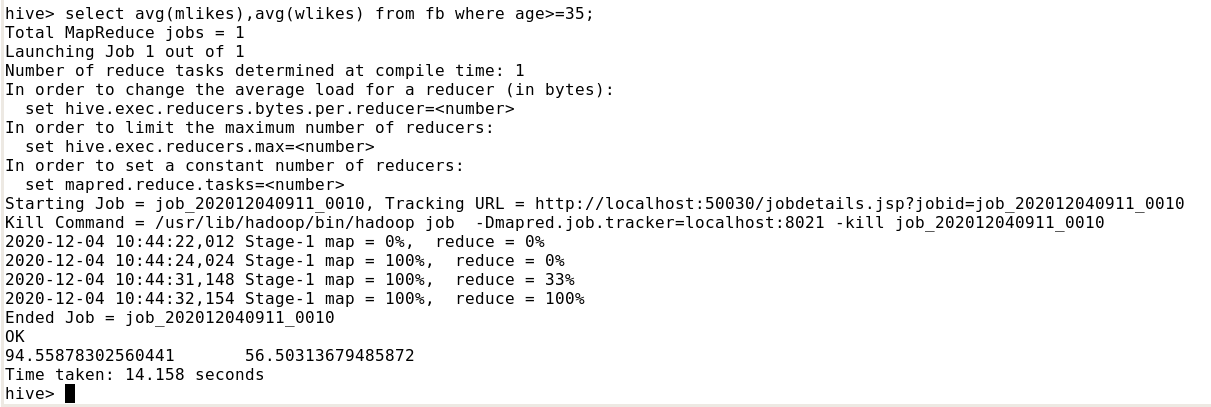
**PROBLEM STATEMENT 5:Find out the count of Facebook users for each birthday month.**

From the above result we can say that most number of users created their account in the month of January so the best time for brand promotion can be considered as January .

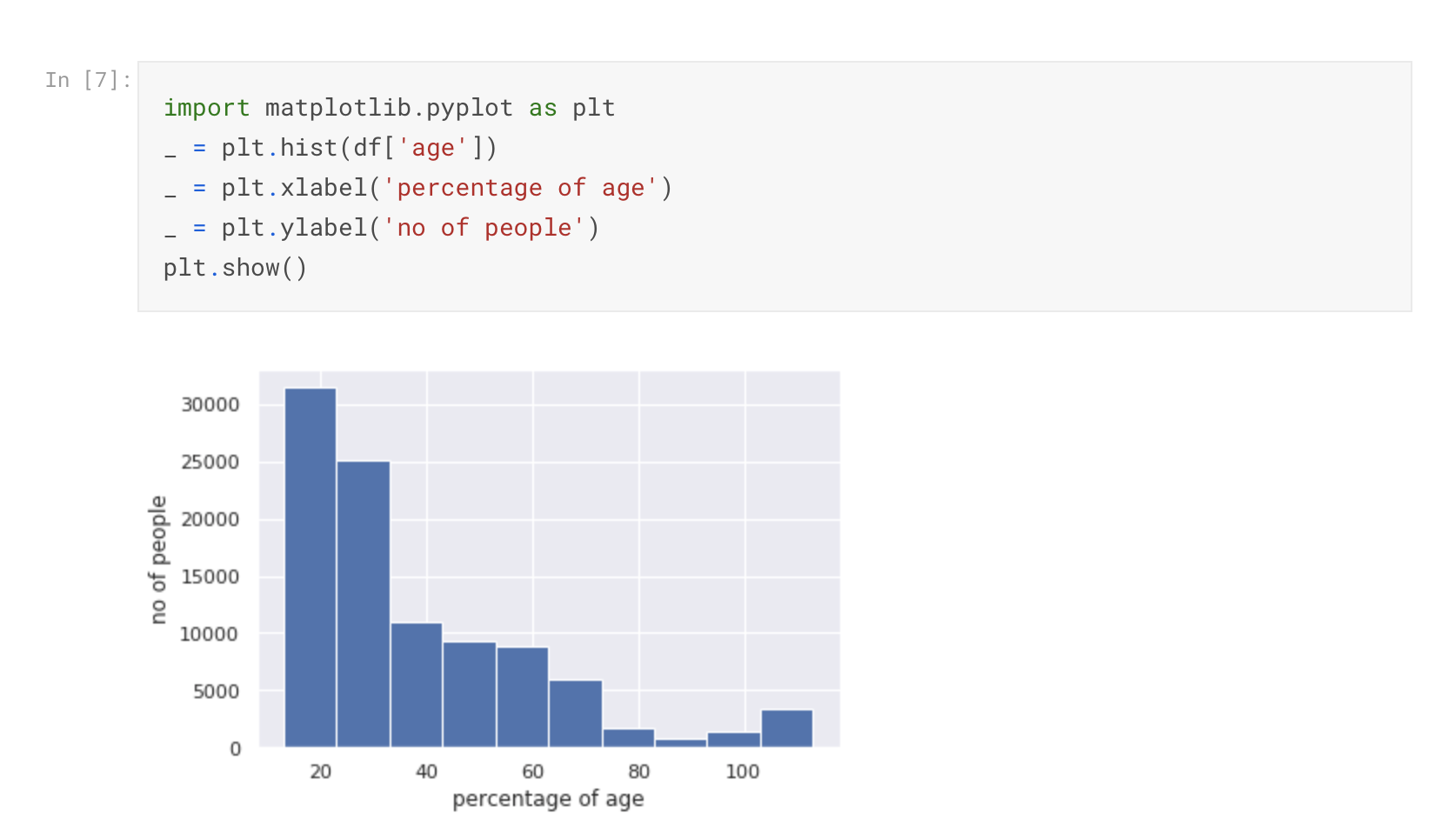
**PROBLEM STATEMENT 6.Do young members use mobile phones or computers for Facebook browsing?**

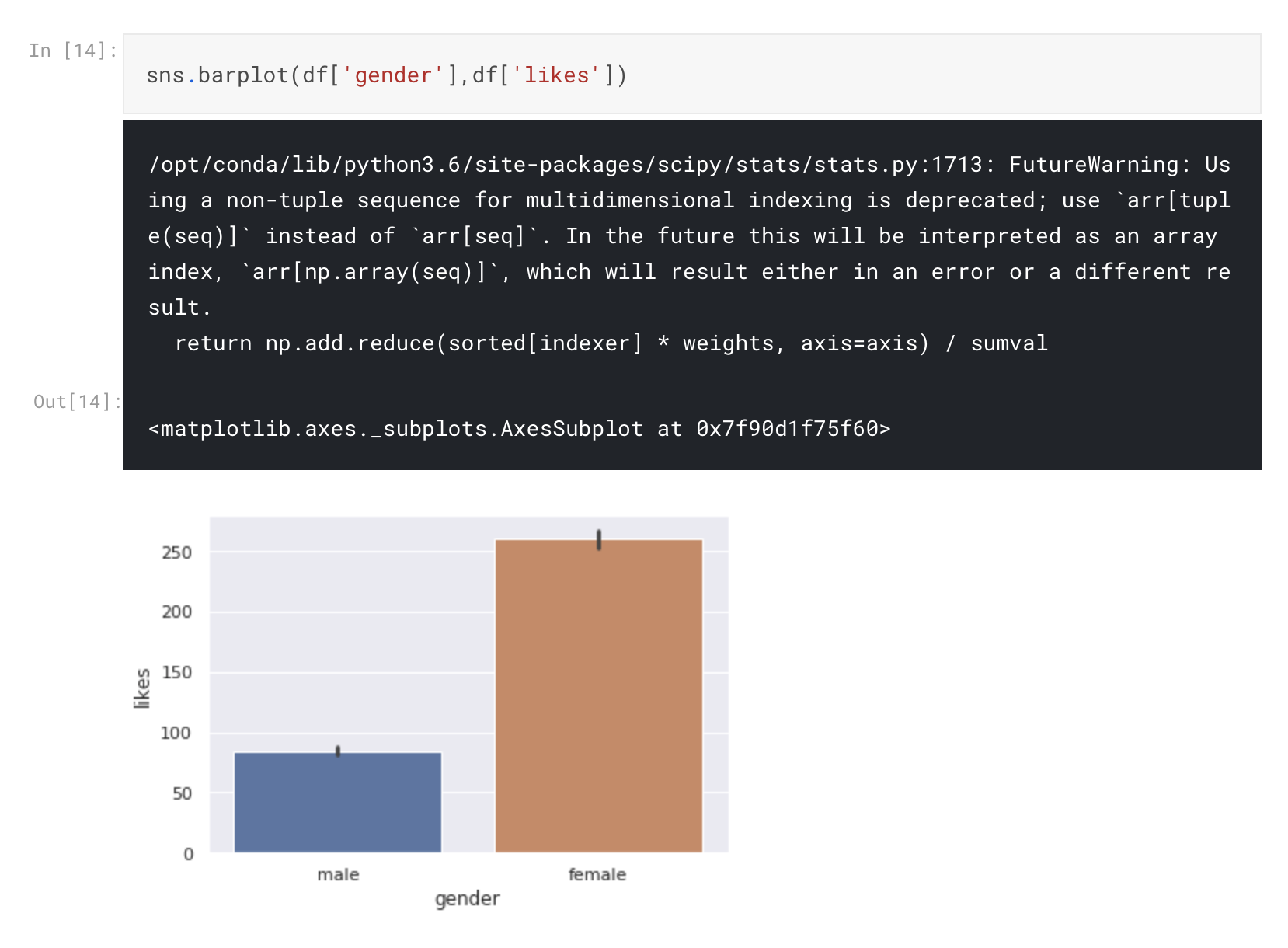


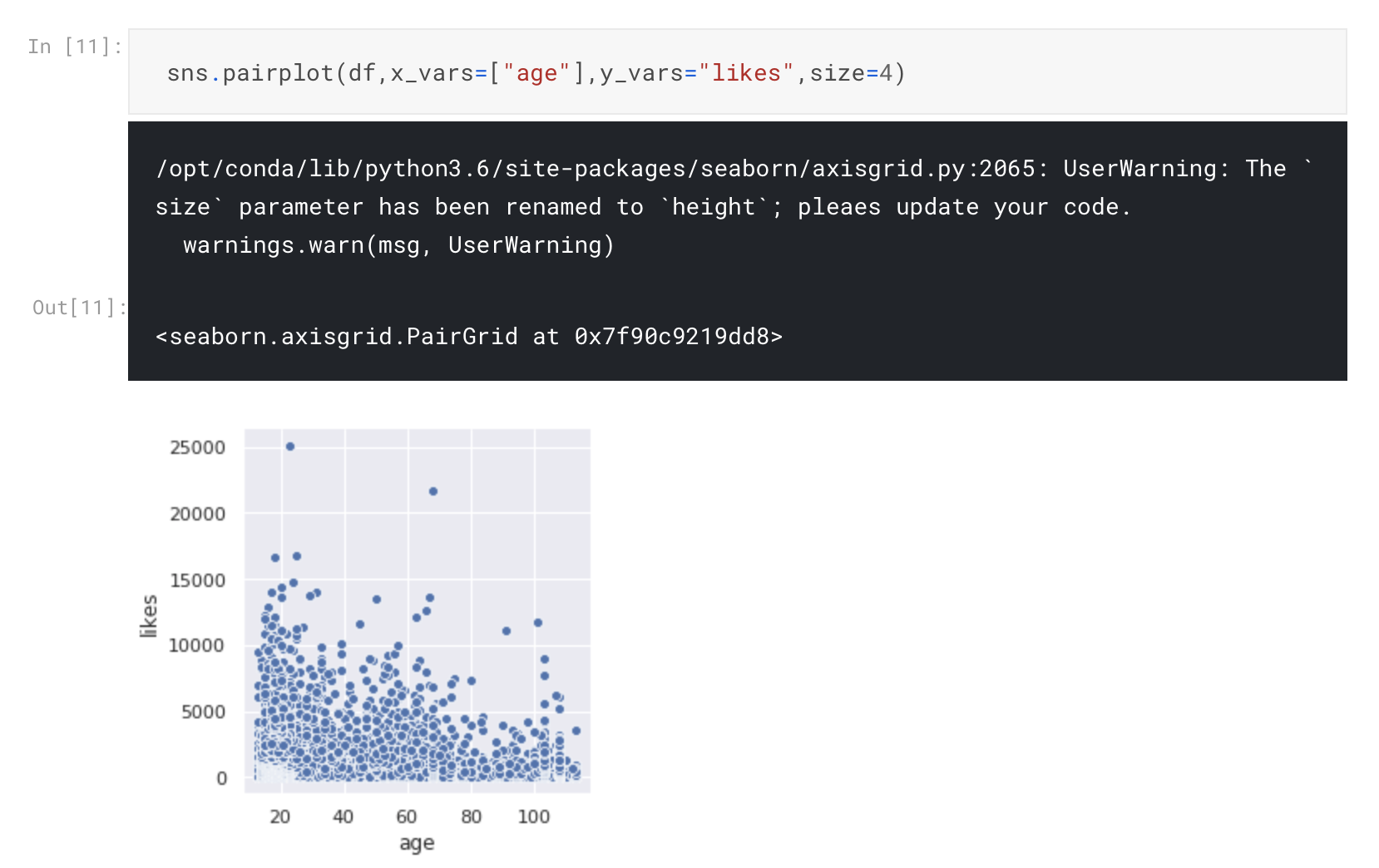
Thus, it can be seen that young members use mobile phones for using Facebook instead of using computers so we can display our ads on mobile phones.

**PROBLEM STATEMENT 7:Do adult members use mobile phones or computers for Facebook browsing?**

I thought that adult users that is above the age of 35 mostly prefer computers for Facebook use, but the result is shocking as we can say adult also prefer mobile phones for use of Facebook but the number of adult users are less then young users .

**PROBLEM STATEMENT 8:Visualisation graph for the age wise number of people on Facebook :**

**PROBLEM STATEMENT 9:Visualisation for the number of likes which was received by male and female:**

**PROBLEM STATEMENT 10:Visualisation for the likes received for the age of the users (Male or Female):**

**image2.png**