Hadoop: Hive Assignment - 2

Problem Statement:

You work for Spring Analytics Pvt. Ltd., a company that specializes in Big Data Analytics. Your organization's latest client is a well-known social media enterprise. Based on the immense amount of data generated in the company, it is planning a targeted advertising campaign for its benefactors. Your firm is provided with this dataset, and it has put forth a set of queries based on that dataset. Use your data warehousing expertise in Apache Hive to help your client with its use case.

Dataset:

https://drive.google.com/file/d/1BzPFG0_vdLtF_bpApo9cssMKliqAyqQe/view?pli=1

Instruction:

You have to use the Facebook Data folder to get the .csv dataset.

Fields In The Dataset:

The following list represents the column names and their descriptions:
- userid: The unique identification number of an account
- age
- dob_day
- dob_year
- dob_month
- gender
- tenure: The number of days since the account has been created
- friend_count: The current number of friends added to the account
- friendships_initiated: The total friend requests sent
- likes: The number of likes done by the user
- likes_received: The number of likes received by the user
- mobile_likes: The amount of mobile application likes done by the user
- mobile_likes_received: The number of likes received by the user from other mobile users
- www_likes: The amount of web browser likes done by the user
- www_likes_received: The amount of web browser likes received by the user

Tasks To Be Performed:

1. What is the average age of a social media account user?
2. Does the social media platform have a higher number of male users, female users, or gender undisclosed users?
3. In male users, on average, does the age demographic of 13–25 have more or less friends than the demographic of 26–50? Assess this with appropriate statistical reasoning.
4. In female users, on average, does the age demographic of 13–25 have more or less friends than the demographic of 26–50? Assess this with appropriate statistical reasoning.
5. Which gender is more likely to send out a higher number of friend requests on average?
6. With the age demographics of 13–25, 26–35, and 36–50 as focal points, evaluate the comparison between mobile application usage and web browser usage when accessing the social media website. Use this to determine if mobile phones have indeed taken over the digital marketspace.