Instagram User Analytics

SQL Fundamentals

Description:

The goal of this project is to use SQL queries to analyse various parts of Instagram's data. The project covers marketing campaign and investor metrics duties. We will examine the given database to answer questions about rewarding loyal users, engaging inactive users, announcing contest winners, researching hashtags, initiating ad campaigns, analysing user engagement, and identifying phoney accounts.

Approach:

- Database Installation: Import the supplied database into a SQL environment.
- Data Exploration: Analyse the tables to determine their structure and linkages.
- Query Execution: Create SQL queries to get the information needed for each activity.
- Data Analysis: Interpret the results of the queries to get useful insights.

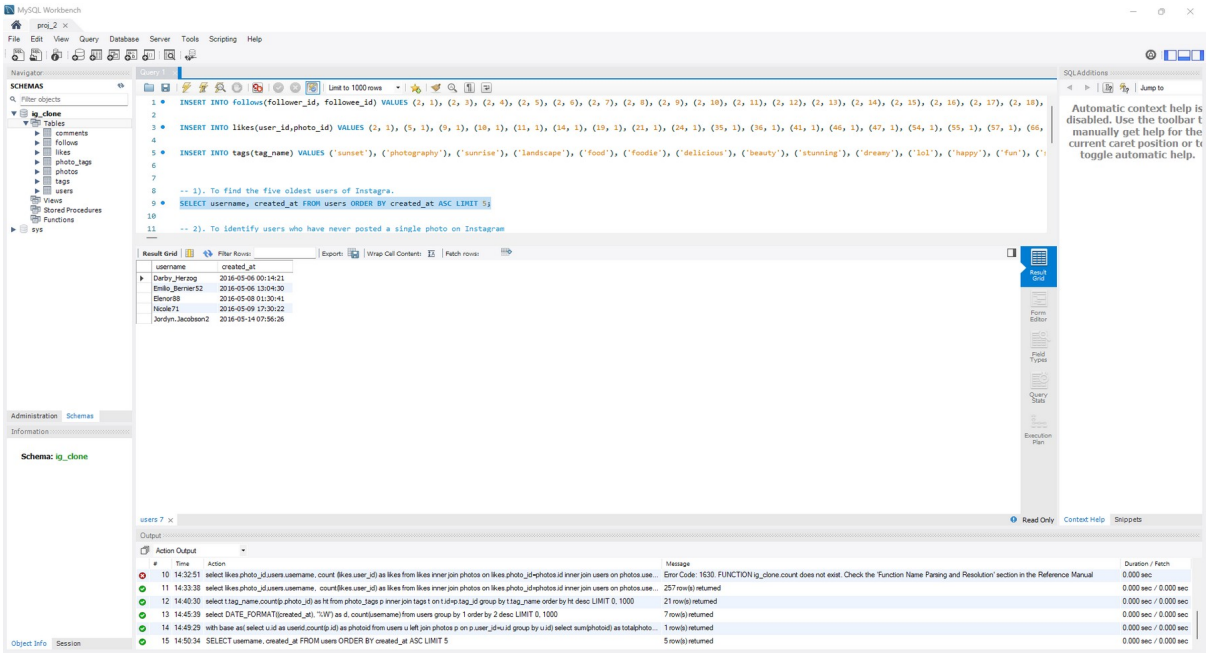
Tech-Stack Used:
For this project, we have utilized SQL environment which is MySQL Workbench

Tasks A:

Marketing: The marketing team wants to launch some campaigns, and they need your help with the following

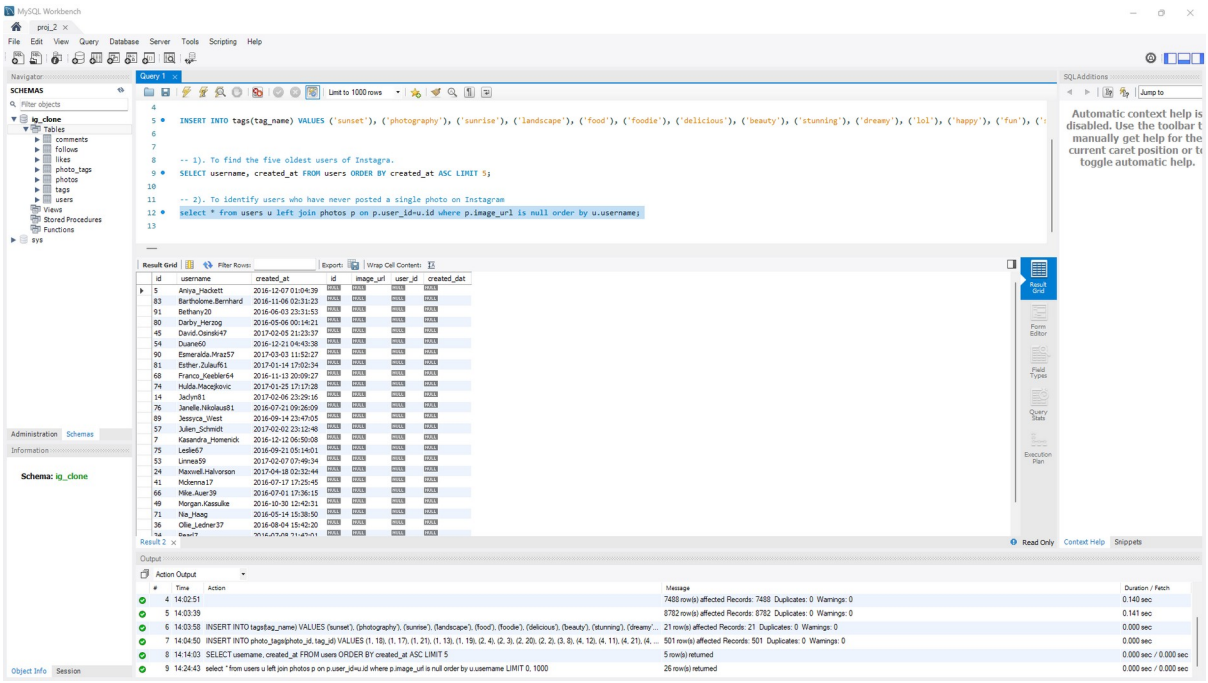
1. **Rewarding Most Loyal Users:** People who have been using the platform for the longest time.
   Your Task: Find the 5 oldest users of the Instagram from the database provided

   Solution: SELECT username, created_at FROM users ORDER BY created_at ASC LIMIT 5;



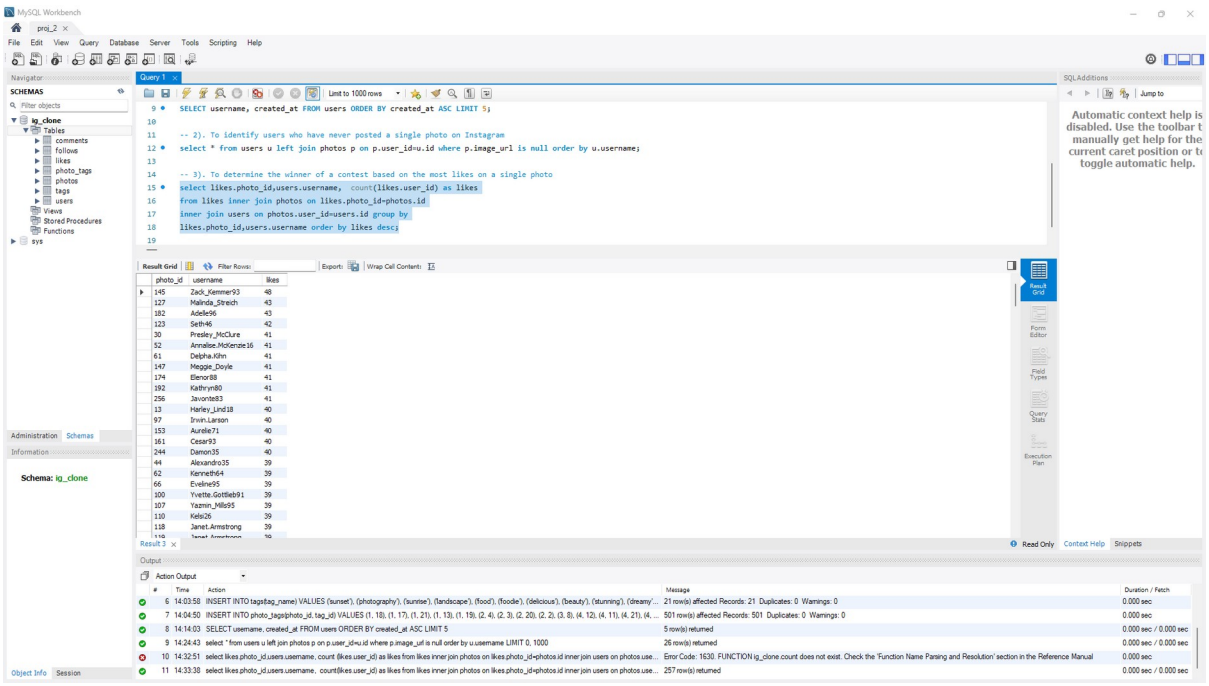
2. **Remind Inactive Users to Start Posting:** By sending them promotional emails to post their 1st photo.
   Your Task: Find the users who have never posted a single photo on Instagram.

   Solution: select * from users u left join photos p on p.user_id=u.id where p.image_url is null order by u.username;



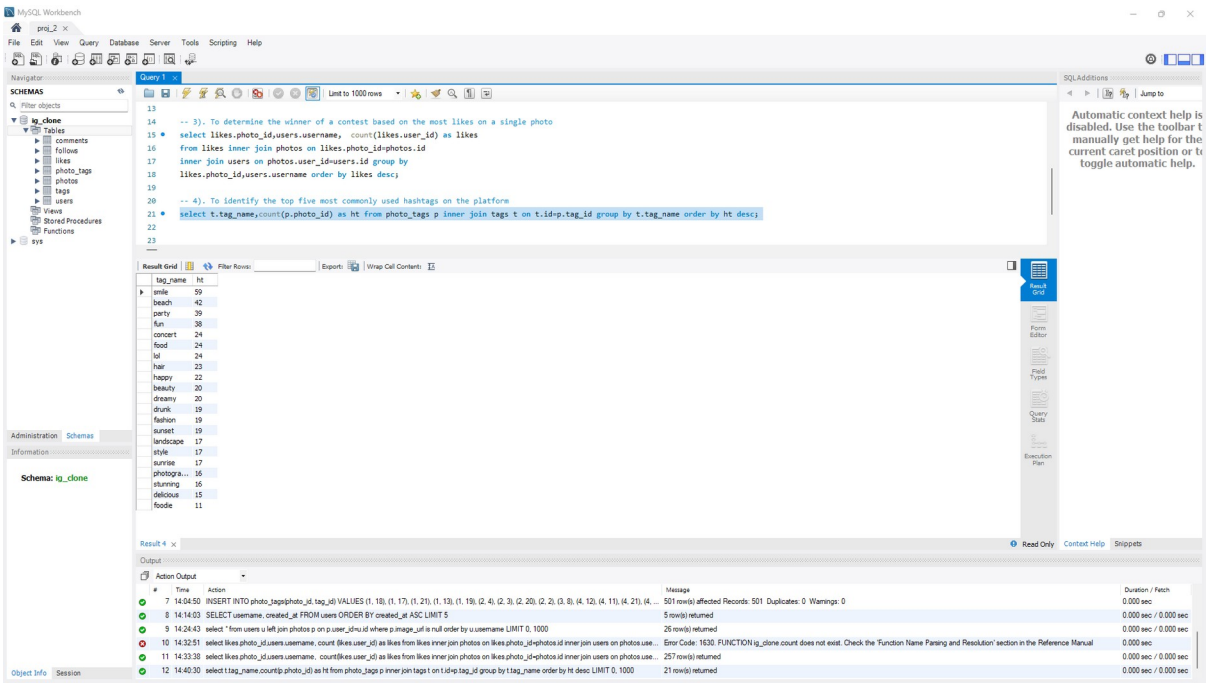
3. **Declaring Contest Winner:** The team started a contest and the user who gets the most likes on a single photo will win the contest now they wish to declare the winner.
   Your Task: Identify the winner of the contest and provide their details to the team.

Solution: select likes.photo_id,users.username, count(likes.user_id) as likes from likes inner join photos on likes.photo_id=photos.id inner join users on photos.user_id=users.id group by likes.photo_id,users.username order by likes desc;



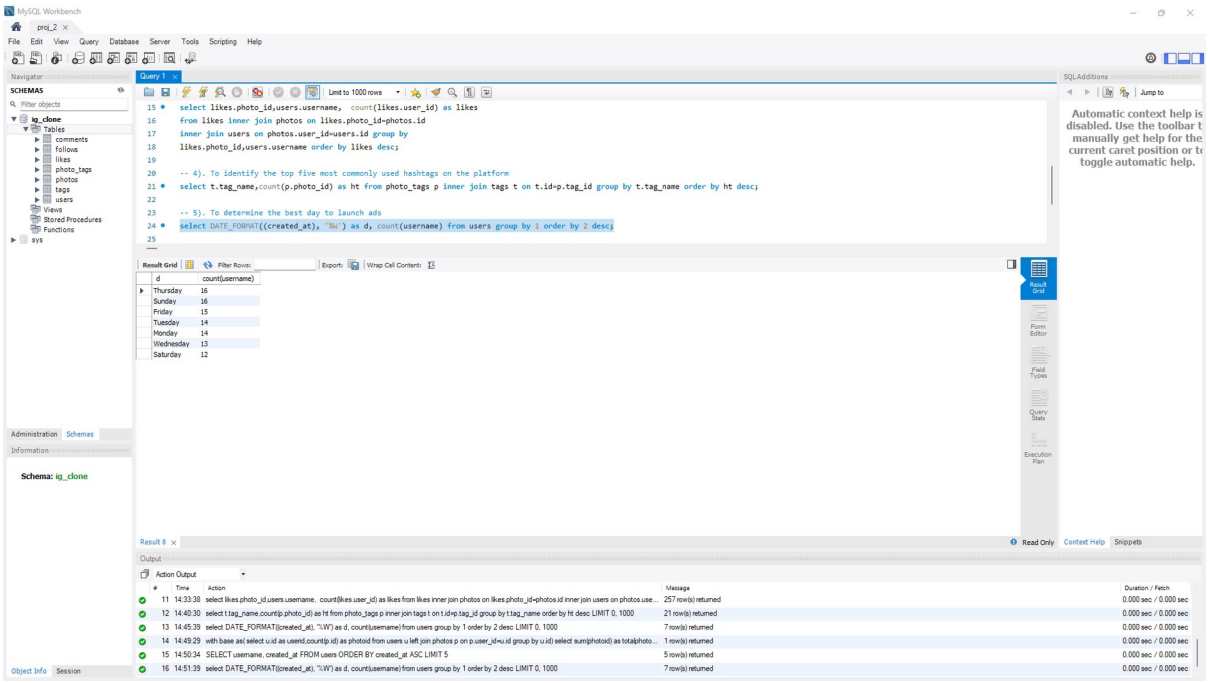
4. **Hashtag Researching:** A partner brand wants to know, which hashtags to use in the post to reach the most people on the platform.
Your Task: Identify and suggest the top 5 most commonly used hashtags on the platform.

Solution: select t.tag_name,count(p.photo_id) as ht from photo_tags p inner join tags t on t.id=p.tag_id group by t.tag_name order by ht desc;



5. **Launch AD Campaign:** The team wants to know, which day would be the best day to launch ADs.
   Your Task: What day of the week do most users register on? Provide insights on when to schedule an ad campaign

   Solution: select DATE_FORMAT((created_at), '%W') as d, count(username) from users group by 1 order by 2 desc;



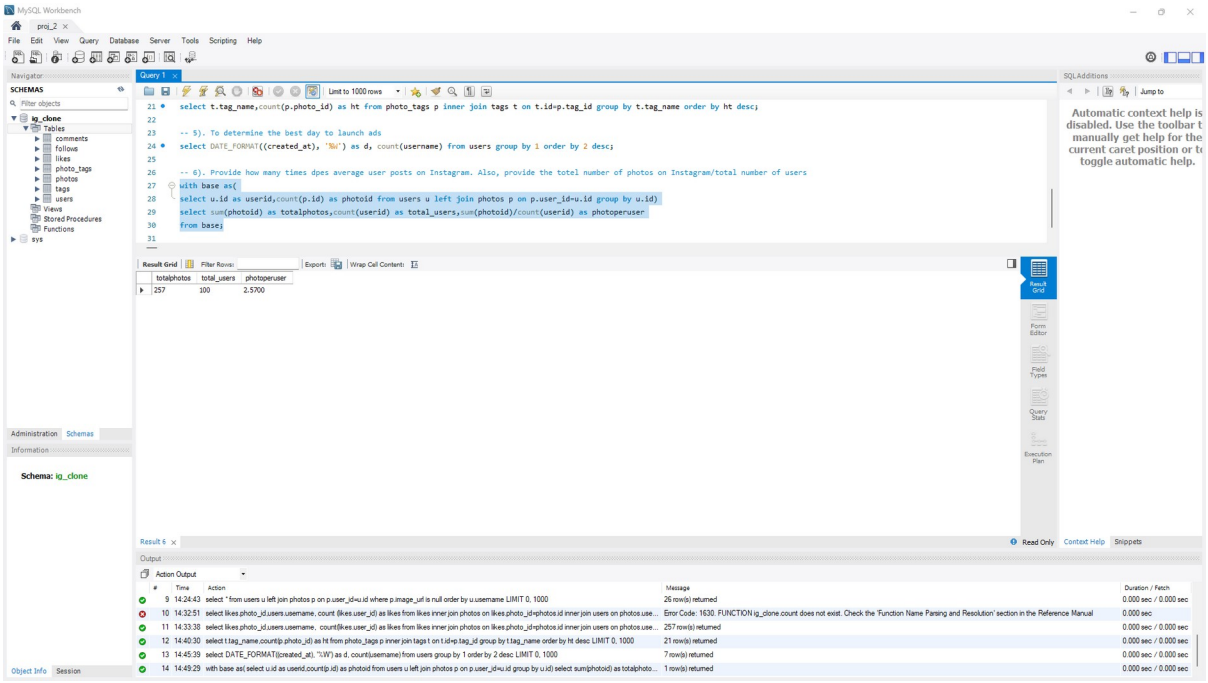
Tasks B:

Investor Metrics: Our investors want to know if Instagram is performing well and is not becoming redundant like Facebook, they want to assess the app on the following grounds

1. **User Engagement:** Are users still as active and post on Instagram or they are making fewer posts
   Your Task: Provide how many times does average user posts on Instagram. Also, provide the total number of photos on Instagram/total number of users.

   Solution: with base as(select u.id as userid,count(p.id) as photoid from users u left join photos p on p.user_id=u.id group by u.id)

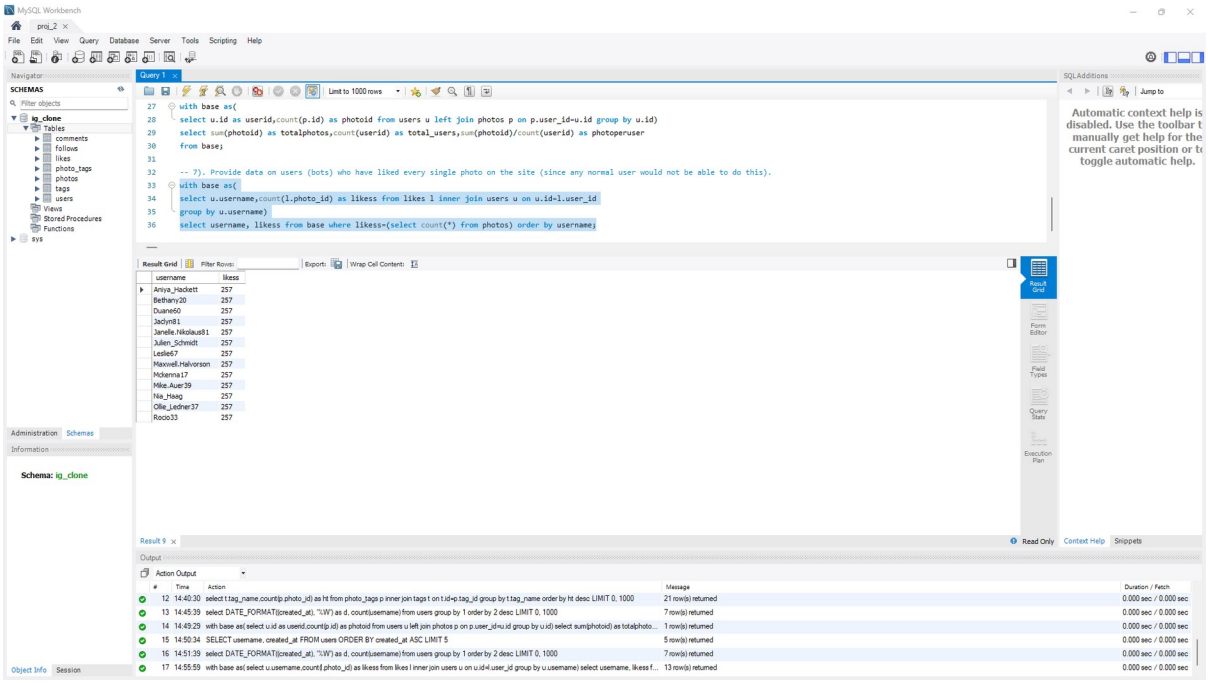
   select sum(photoid) as totalphotos,count(userid) as total_users,sum(photoid)/count(userid) as photoperuser from base;



2. **Bots & Fake Accounts:** The investors want to know if the platform is crowded with fake and dummy accounts
Your Task: Provide data on users (bots) who have liked every single photo on the site (since any normal user would not be able to do this).

Solution: with base as(select u.username,count(l.photo_id) as likess from likes l inner join users u on u.id=l.user_id group by u.username)

select username, likess from base where likess=(select count(*) from photos) order by username;



Results:

As a result of finishing this project, we were able to:

- Determine the five oldest Instagram users.
- Determine which individuals have never submitted a single photo.
- The most likes on a single photo determines the winner of a contest.
- Please suggest the top five most popular hashtags on the platform.
- Based on user registration habits, provide information about the ideal day to launch ADs.
- Determine the average number of posts made per user and evaluate user involvement.
- Identify people who have liked each and every photo, which might indicate the presence of bots or fraudulent accounts.