Project 2:
Instagram User Analytics

Kiruba Shankar S

Data Analytics Trainee

Description:

This project aims to analyze Instagram user data to provide valuable insights for the marketing and product teams, by using MySQL Workbench to extract the date from the database. Team can use this data to improve business, improve user engagement by identifying activities of the users and can identify bots and fake accounts
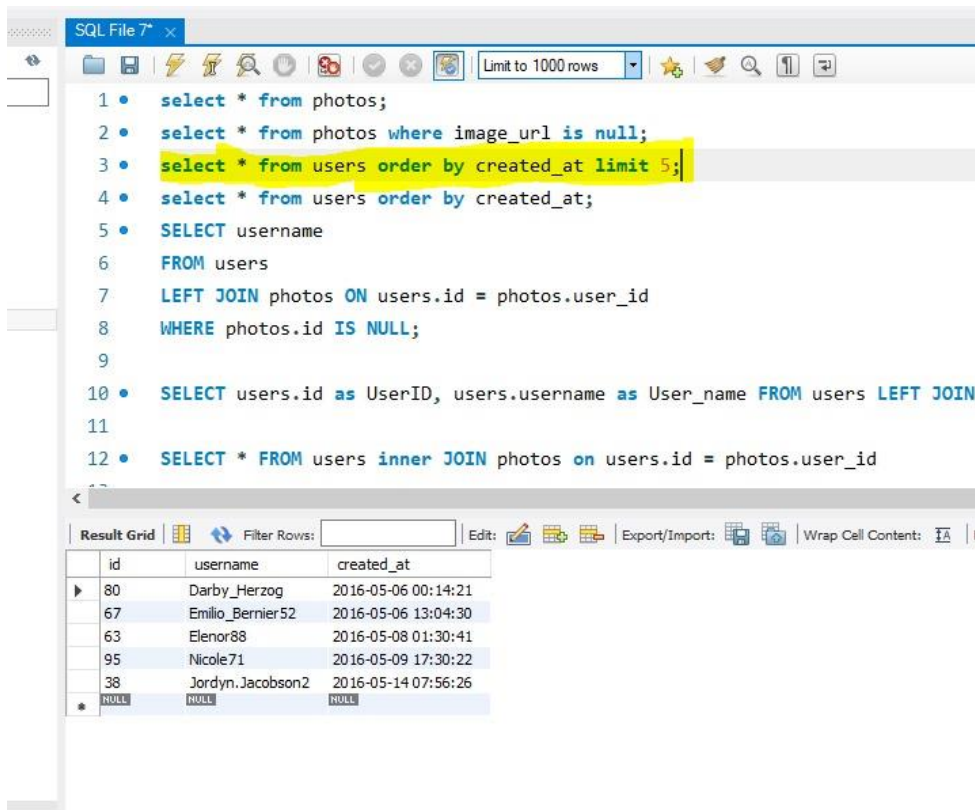
SQL Tasks

A) Marketing Analysis:

1: Loyal User Reward:

The marketing team wants to reward the most loyal users, i.e., those who have been using the platform for the longest time.
Task: Identify the five oldest users on Instagram from the provided database.

Query:

select * from users order by created_at limit 5;

Insight:

These are the top five users who have been using Instagram for the longest time. Rewarding these loyal users can enhance customer satisfaction and promote long-term engagement.

2: Inactive User Engagement:

The team wants to encourage inactive users to start posting by sending them promotional emails.
Task: Identify users who have never posted a single photo on Instagram.

```
10 •   SELECT users.id as UserID, users.username as User_name FROM users LEFT JOIN photos on users.id = photos.user_id where photos.id is null;
11
12 •   SELECT * FROM users inner JOIN photos on users.id = photos.user_id
```

UserID	User_name
5	Aniya_Hackett
7	Kasandra_Homenick
14	Jaclyn81
21	Rocio33
24	Maxwell.Halvorson
25	Tierra.Trantow
34	Pearl7
36	Ollie_Ledner37
41	Mckenna17
45	David.Osinski47
49	Morgan.Kassulke
53	Linnea59
54	Duane60
57	Julien_Schmidt
66	Mike.Auer39
68	Franco_Keebler64
71	Nia_Haag
74	Hulda.Macejkovic
75	Leslie67
76	Janelle.Nikolaus81
80	Darby_Herzog
81	Esther.Zulauf61
83	Bartholome.Bernhard
89	Jessyca_West
90	Esmeralda.Mraz57
91	Bethany20

Query:

 SELECT users.id as UserID, users.username as User_name FROM users LEFT JOIN photos on users.id = photos.user_id where photos.id is null;
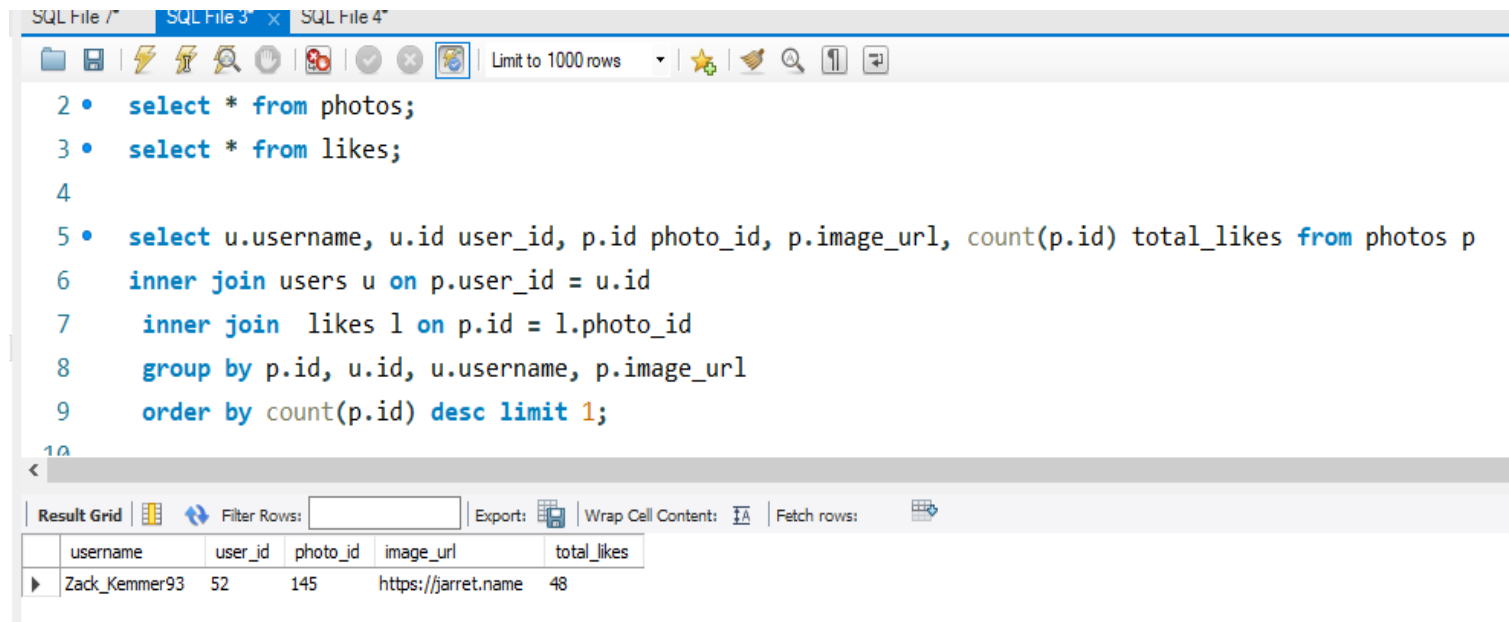
Insight:

These are the users who have not posted a single photo on Instagram, Sending promotional emails to these users can encourage them to start posting and increase overall platform engagement.

3: Contest Winner Declaration:

The team has organized a contest where the user with the most likes on a single photo wins.
Task: Determine the winner of the contest and provide their details to the team.

Query:

select u.username, u.id user_id, p.id photo_id, p.image_url, count(p.id) total_likes
from photos p
inner join users u on p.user_id = u.id
 inner join likes l on p.id = l.photo_id group by p.id, u.id, u.username, p.image_url
order by count(p.id) desc limit 1;

Insight:

The user Zack_Kemmer93 (user ID 52) Is the one who got more likes for a single photo. This user can be declared as the contest winner and awarded accordingly. Recognizing active and popular users can boost engagement and participation in future contests.

4: Hashtag Research:

A partner brand wants to know the most popular hashtags to use in their posts to reach the most people.
Task: Identify and suggest the top five most commonly used hashtags on the platform.

```
15
16 •    select t.id, t.tag_name, count(t.id) as most_used
17      from tags t
18      inner join photo_tags pt on t.id = pt.tag_id
19      group by t.id, t.tag_name
20      order by count(t.id) desc limit 5;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

id	tag_name	most_used
21	smile	59
20	beach	42
17	party	39
13	fun	38
18	concert	24

Query:

select t.id, t.tag_name, count(t.id) as most_used
 from tags t
inner join photo_tags pt on t.id = pt.tag_id
group by t.id, t.tag_name
order by count(t.id) desc limit 5;
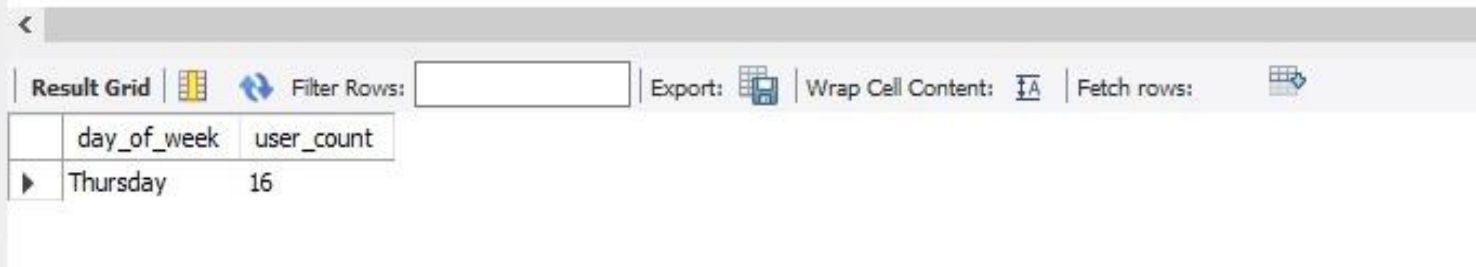
Insight:

These are the popular hashtags on the platform. Partner brands can use these hashtags to reach a wider audience. Using relevant and popular hashtags can significantly increase post visibility and more engagements.

5: Ad Campaign Launch:

The team wants to know the best day of the week to launch ads.

Your Task: Determine the day of the week when most users register on Instagram.
Provide insights on when to schedule an ad campaign.

```
27
28 •   SELECT DAYNAME(created_at) AS day_of_week, COUNT(id) AS user_count
29     FROM users
30     GROUP BY day_of_week
31     ORDER BY user_count DESC
32     LIMIT 1;
33
```

	day_of_week	user_count
▶	Thursday	16

Query:

SELECT DAYNAME(created_at) AS day_of_week, COUNT(id) AS user_count
FROM users
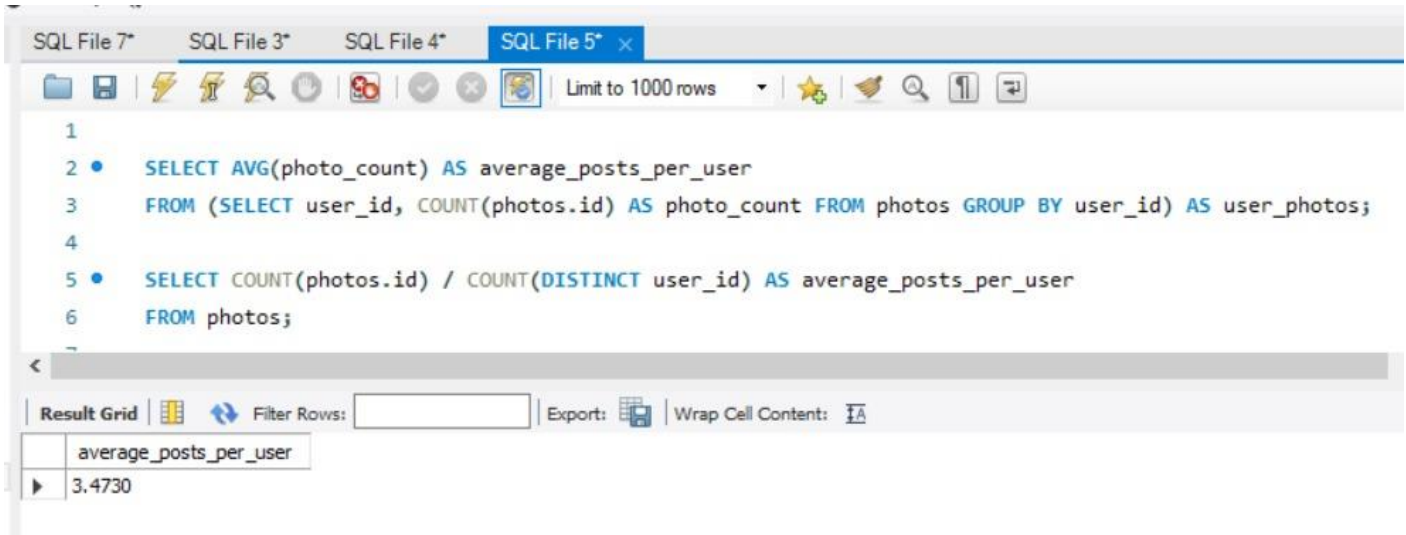GROUP BY day_of_week
ORDER BY user_count DESC
LIMIT 1;

Insight:

The day with the highest number of user registrations is Thursday, which the optimal
day for launching ad campaigns. Scheduling ads on this day can maximize reach and
effectiveness.

B) Investor Metrics:

1: User Engagement:

Investors want to know if users are still active and posting on Instagram or if they are making fewer posts.
Task: Calculate the average number of posts per user on Instagram. Also, provide the total number of photos on Instagram divided by the total number of users.

Query 1:

```
SELECT AVG(photo_count) AS average_posts_per_user
FROM (SELECT user_id, COUNT(photos.id) AS photo_count FROM photos GROUP
BY user_id) AS user_photos;
```

Query 2:

```
SELECT COUNT(photos.id) / COUNT(DISTINCT user_id) AS
average_posts_per_user
FROM photos;
```
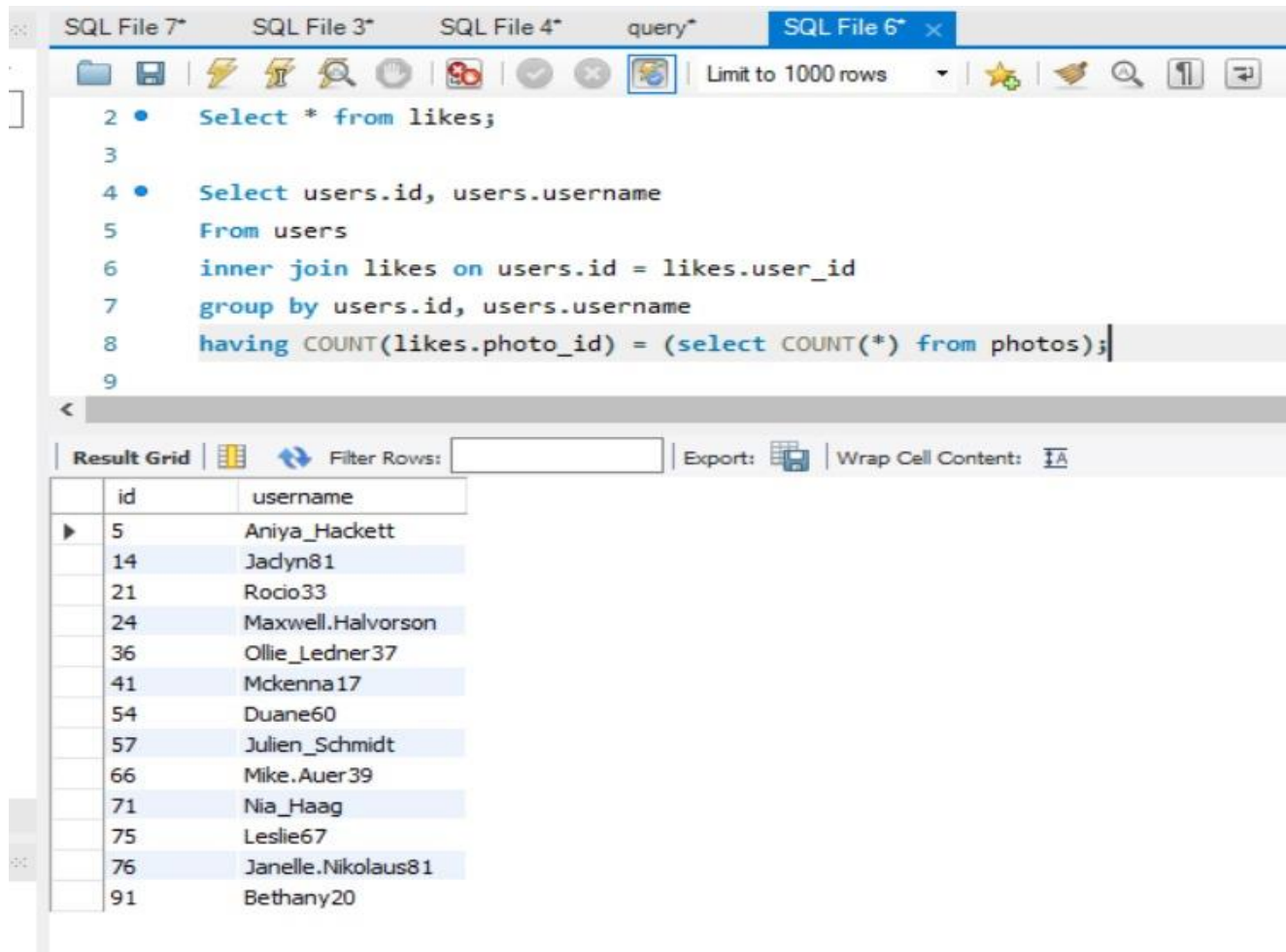
Explanation: The query 1 and 2, which is given above gives the same output 3.4730 which is the average number of posts per user and total number of photos on Instagram, which is divided by the total number of users.

Insights:
With the given results, we could see that users are still active, but there is room to improve the activity, we can send promotional emails and we can add new features to Instagram which will drive higher user engagement.

2: Bots & Fake Accounts:

Investors want to know if the platform is crowded with fake and dummy accounts.
Task: Identify users (potential bots) who have liked every single photo on the site, as
this is not typically possible for a normal user.

Query:

SELECT users.id, users.username
FROM users
inner JOIN likes ON users.id = likes.user_id
GROUP BY users.id, users.username
HAVING COUNT(likes.photo_id) = (SELECT COUNT(*) FROM photos);

Insights:

These are the users who have liked each and every photo on the platform, hence it is
impossible by normal user to like every single post. These accounts are likely bots or
fake accounts. Identifying and removing such accounts can improve the platform's
authenticity and user trust.

Conclusion:

The overall analysis provided insights to optimize marketing strategies, enhance user engagement, and improve the overall user experience. These insights will help by contributing various information to internal teams by providing them valuable information, which will help them to enhance the loyal Instagram users and also help the team to identify fake accounts and bots, which will improve Instagram authenticity and trust for the users, which will be useful for future development and growth of Instagram.