Instagram User Analytics

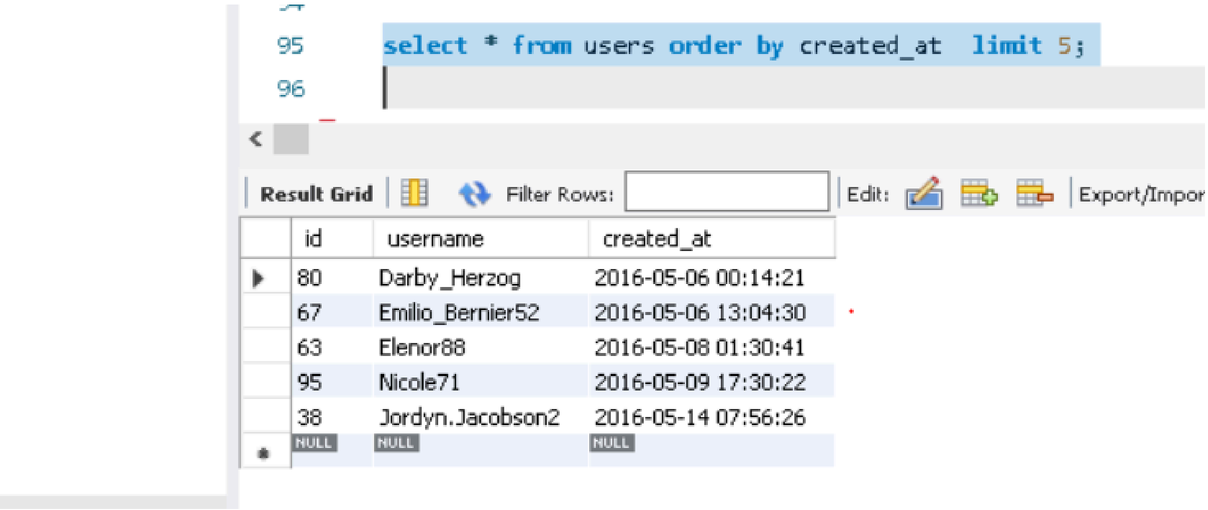
SUMMARY :

This project provides a comprehensive analysis of user engagement and behavior on Instagram. The marketing analysis section helps in identifying key user groups and trends, such as the most loyal users, inactive users, contest winners, popular hashtags, and the best day for launching ad campaigns. The investor metrics section focuses on user engagement and identifying potential fake accounts, providing valuable insights for platform optimization and user retention strategies.

ANSWERS :

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.
   Your Task: Identify the five oldest users on Instagram from the provided database.



```
select * from users order by created_at  limit 5;
```

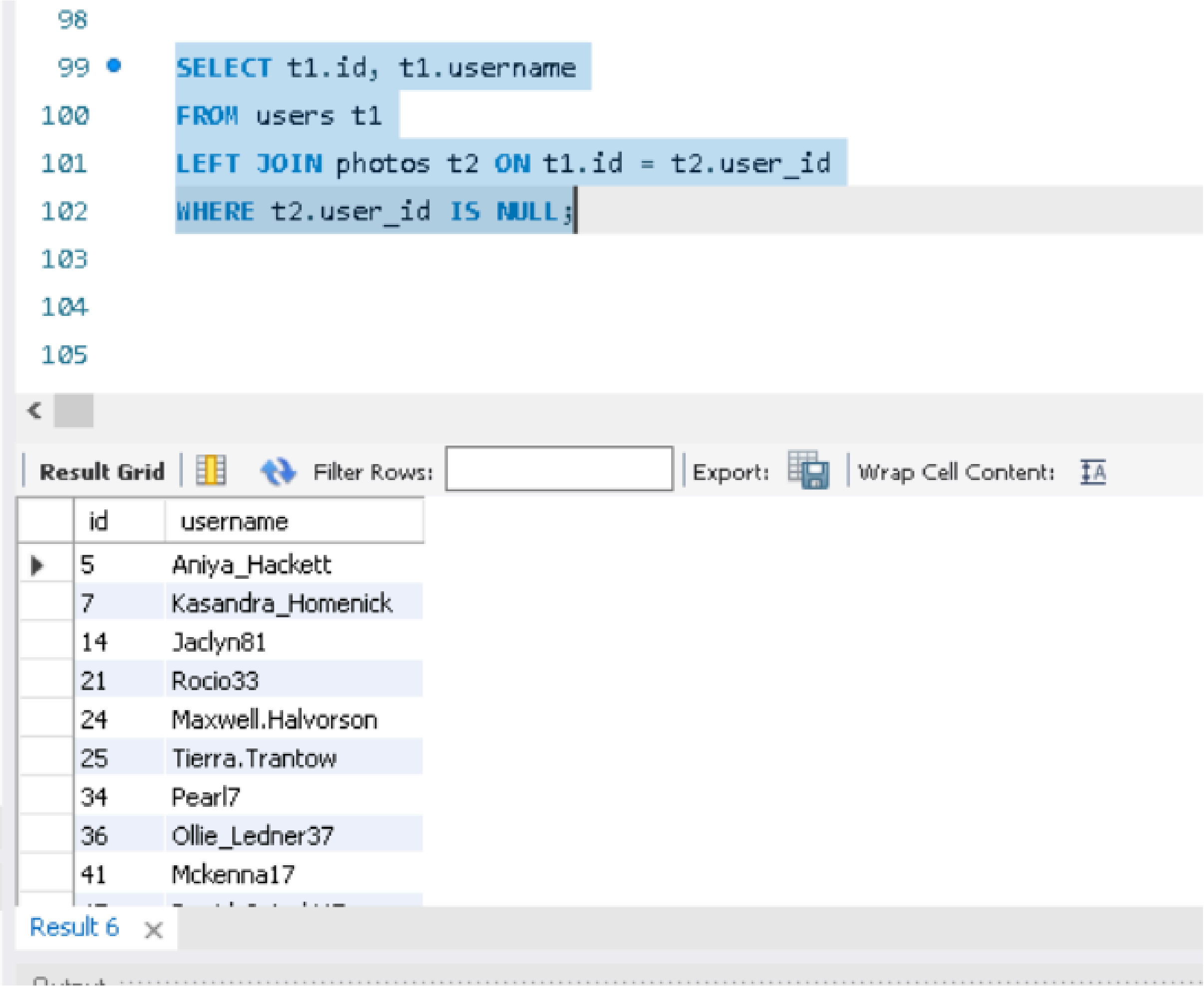
id	username	created_at
80	Darby_Herzog	2016-05-06 00:14:21
67	Emilio_Bernier52	2016-05-06 13:04:30
63	Elenor88	2016-05-08 01:30:41
95	Nicole71	2016-05-09 17:30:22
38	Jordyn.Jacobson2	2016-05-14 07:56:26
NULL	NULL	NULL

Explanation:

ORDER BY created_at ASC: This sorts the users in ascending order based on the created_at timestamp, meaning the oldest users come first.

LIMIT 5: This restricts the result set to the first five rows, which correspond to the five oldest users.

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



```
SELECT t1.id, t1.username
FROM users t1
LEFT JOIN photos t2 ON t1.id = t2.user_id
WHERE t2.user_id IS NULL;
```

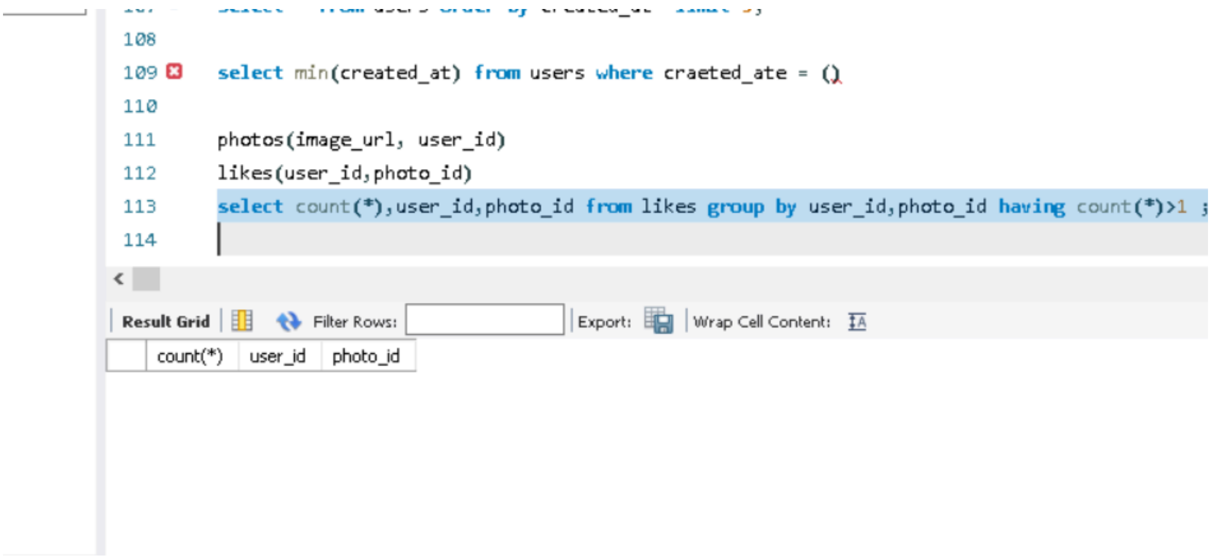
id	username
5	Aniya_Hackett
7	Kasandra_Homenick
14	Jaclyn81
21	Rocio33
24	Maxwell.Halvorson
25	Tierra.Trantow
34	Pearl7
36	Ollie_Ledner37
41	Mckenna17

Explanation:

In user table there will be all the ids of the user and in photos only those user will be there who have added photos so we have to find those user id which is not present in photos ‘s user id

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

user id In this question as mentioned most likes on single photo so in likes table i have photo_id and user_id , at first i tried to find out any duplicates i wanted to search for multiple entries for one single user and with same photo id so that i can count the likes the for a single photo but in likes table i have distinct photo_id only so i have count of one single user with distinct photos

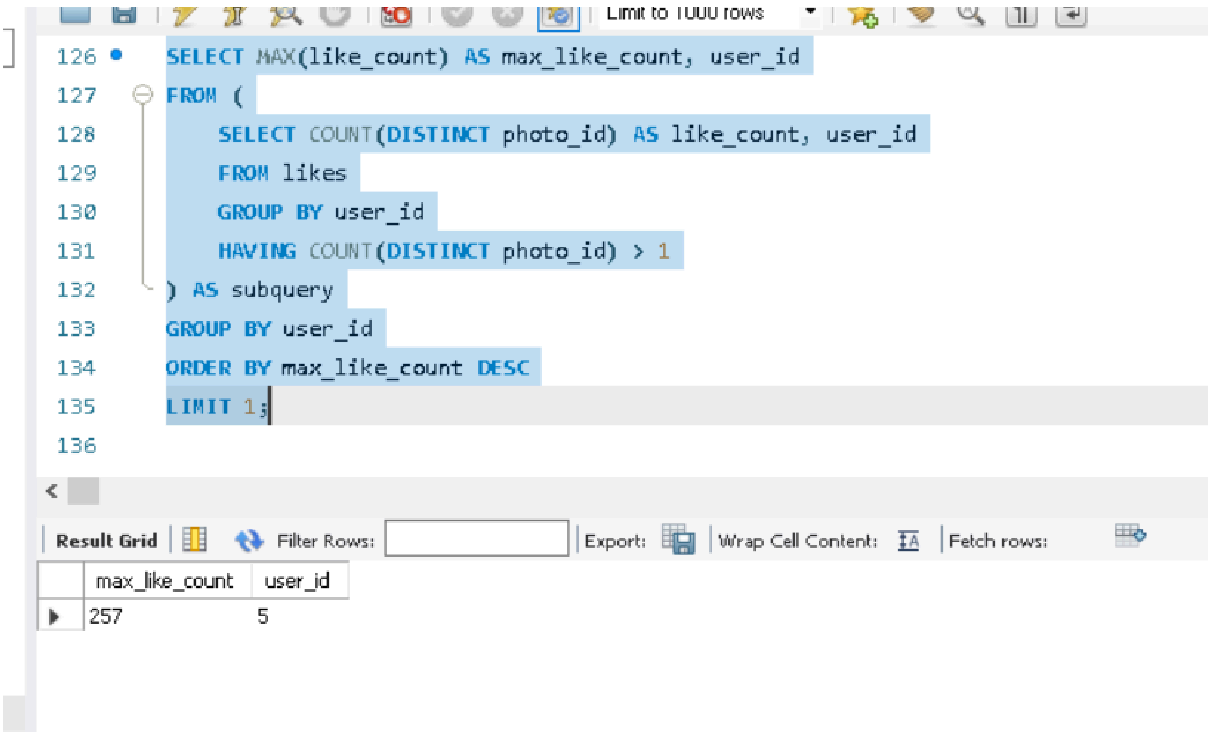


```
select min(created_at) from users where craeted_ate = ()

photos(image_url, user_id)
likes(user_id,photo_id)
select count(*),user_id,photo_id from likes group by user_id,photo_id having count(*)>1 ;
```

count(*)	user_id	photo_id

No duplicates



```
SELECT MAX(like_count) AS max_like_count, user_id
FROM (
    SELECT COUNT(DISTINCT photo_id) AS like_count, user_id
    FROM likes
    GROUP BY user_id
    HAVING COUNT(DISTINCT photo_id) > 1
) AS subquery
GROUP BY user_id
ORDER BY max_like_count DESC
LIMIT 1;
```

max_like_count	user_id
257	5

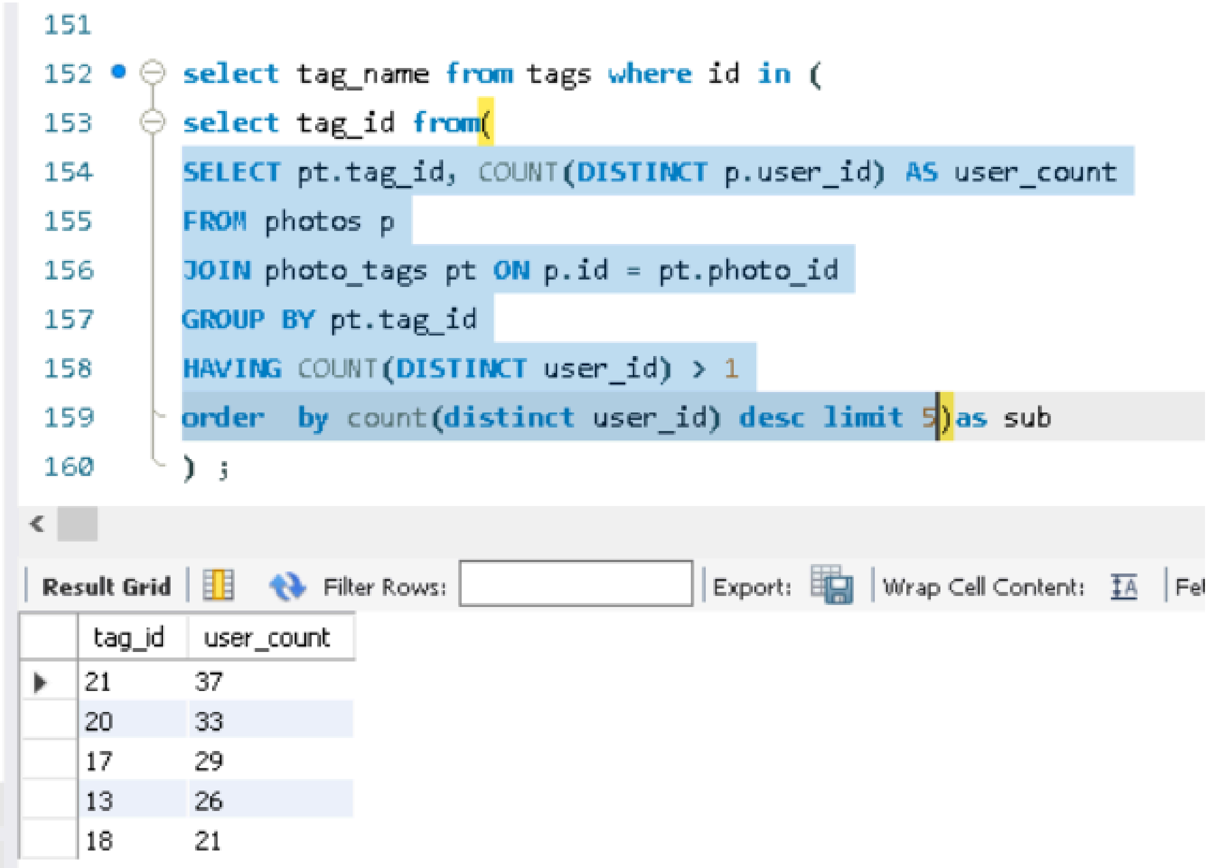
In this subquery i found the distinct photo id with the same user id and counted the the number given alisa as likes_count and in outer query i found the max of likes count with the user_id order by desc max_like_count so max will come first and limit 1 to give one row of max one

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

In this i will try to find the count user used each tags i.e per tag how many users then i will find the top 5 of it and display it

To find the count of user per tag i will use it as subquery and in outer query and limit to max 5 and in outer query will find the tags by using the tag id from inner query

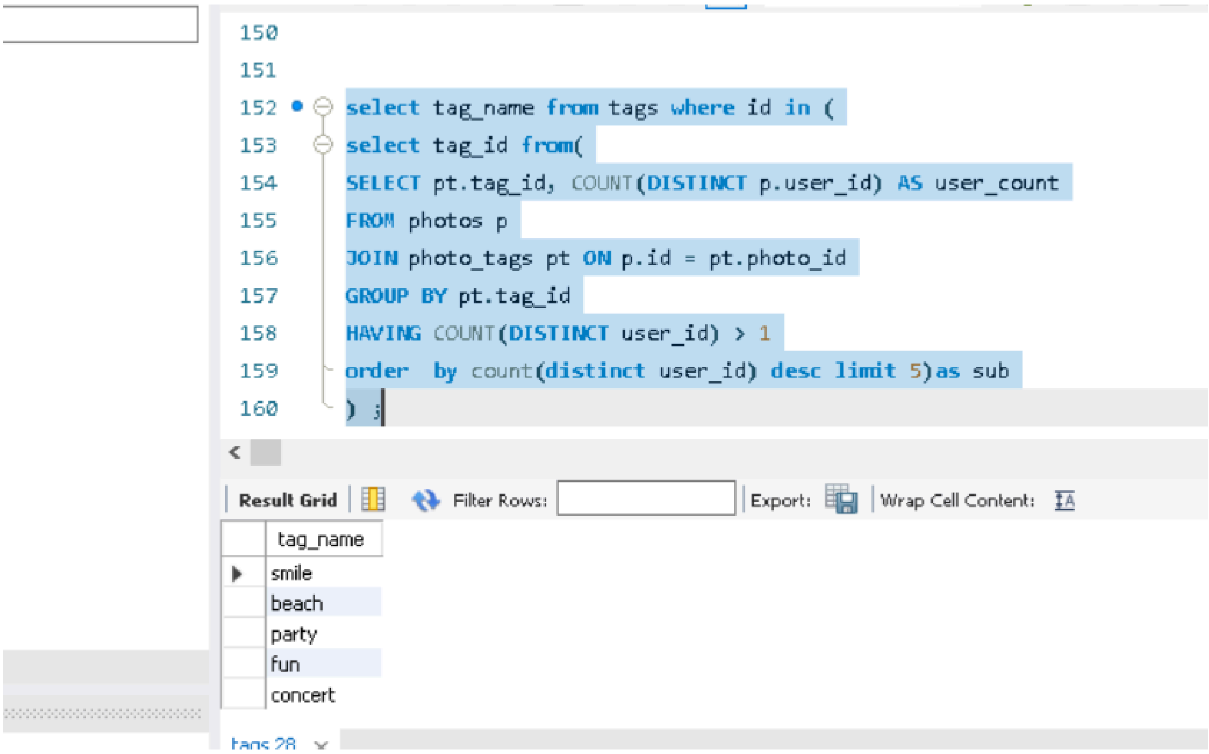
INNER QUERY :



```
select tag_name from tags where id in (
select tag_id from(
SELECT pt.tag_id, COUNT(DISTINCT p.user_id) AS user_count
FROM photos p
JOIN photo_tags pt ON p.id = pt.photo_id
GROUP BY pt.tag_id
HAVING COUNT(DISTINCT user_id) > 1
order  by count(distinct user_id) desc limit 5)as sub
) ;
```

tag_id	user_count
21	37
20	33
17	29
13	26
18	21

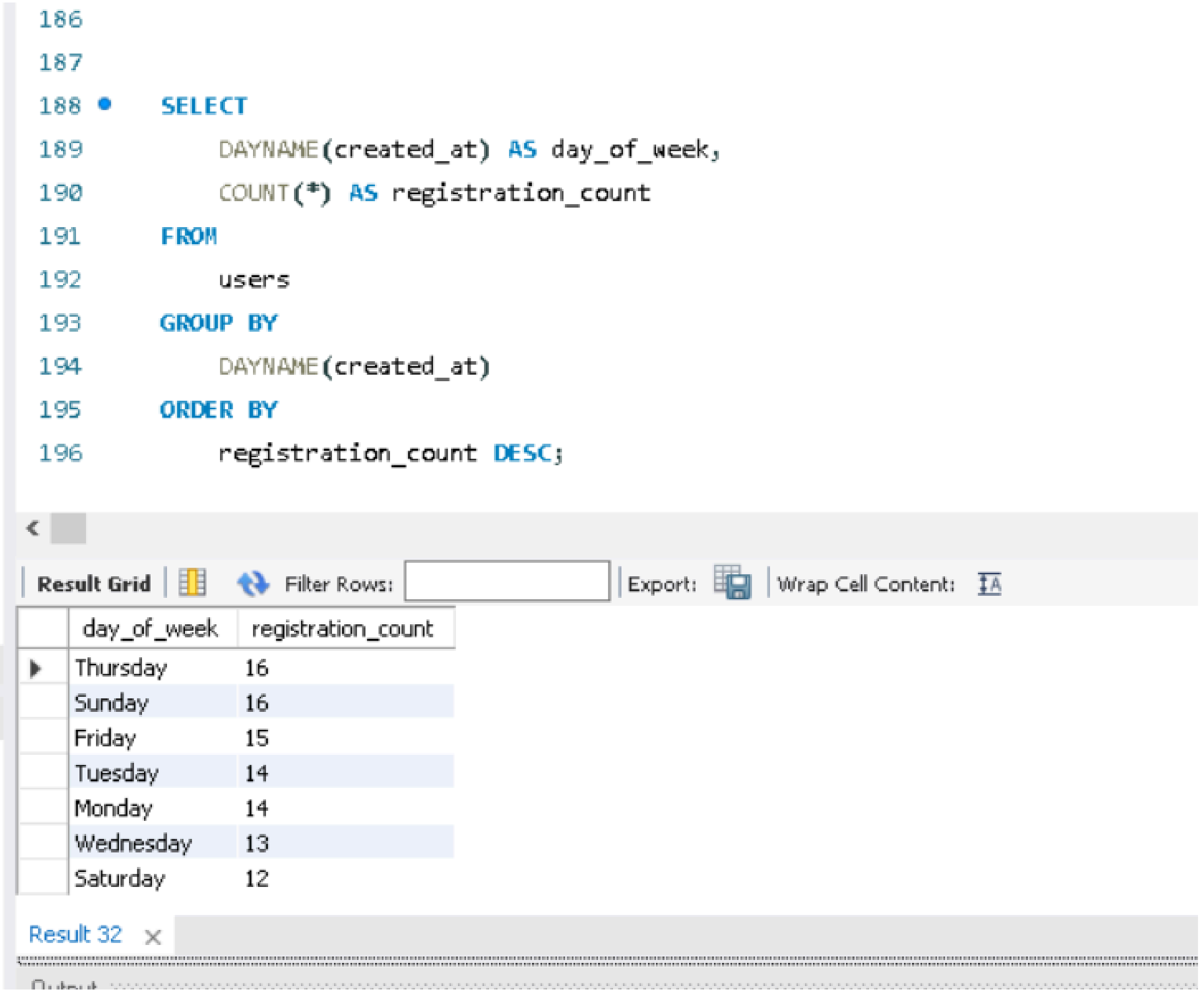
FULL QUERY RESULT :



```
select tag_name from tags where id in (
select tag_id from(
SELECT pt.tag_id, COUNT(DISTINCT p.user_id) AS user_count
FROM photos p
JOIN photo_tags pt ON p.id = pt.photo_id
GROUP BY pt.tag_id
HAVING COUNT(DISTINCT user_id) > 1
order  by count(distinct user_id) desc limit 5)as sub
) ;
```

tag_name
smile
beach
party
fun
concert

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.

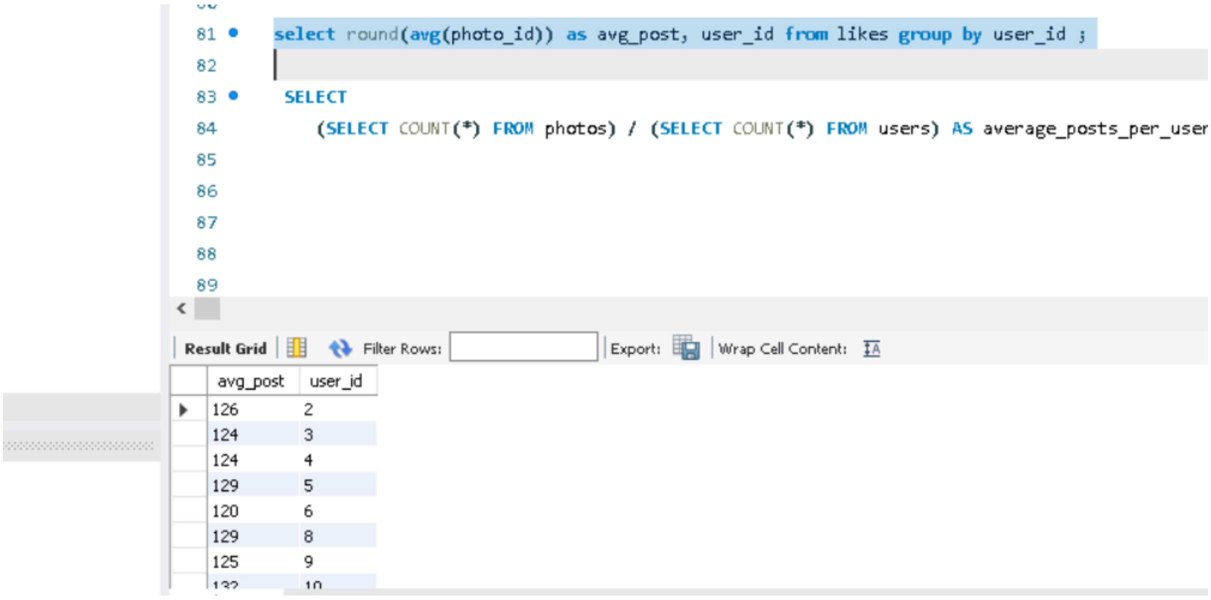


```sql
SELECT
    DAYNAME(created_at) AS day_of_week,
    COUNT(*) AS registration_count
FROM
    users
GROUP BY
    DAYNAME(created_at)
ORDER BY
    registration_count DESC;
```

day_of_week	registration_count
Thursday	16
Sunday	16
Friday	15
Tuesday	14
Monday	14
Wednesday	13
Saturday	12

B) Investor Metrics:

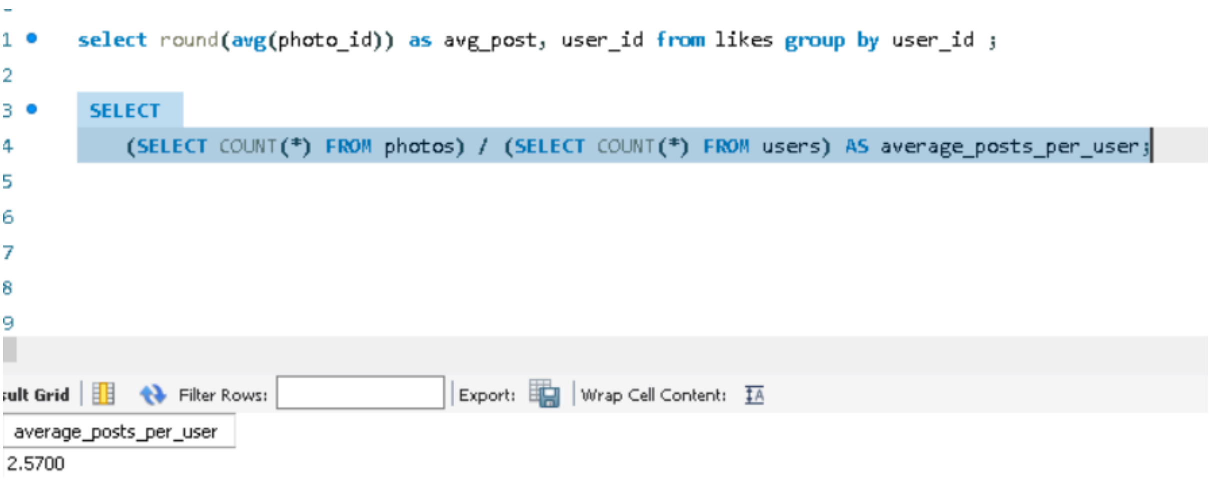
1. User Engagement: Investors want to know if users are still active and posting on Instagram or if they are making fewer posts.
Your Task: Calculate the average number of posts per user on Instagram. Also, provide the total number of photos on Instagram divided by the total number of users.



```
select round(avg(photo_id)) as avg_post, user_id from likes group by user_id ;

SELECT
    (SELECT COUNT(*) FROM photos) / (SELECT COUNT(*) FROM users) AS average_posts_per_user
```

avg_post	user_id
126	2
124	3
124	4
129	5
120	6
129	8
125	9
132	10



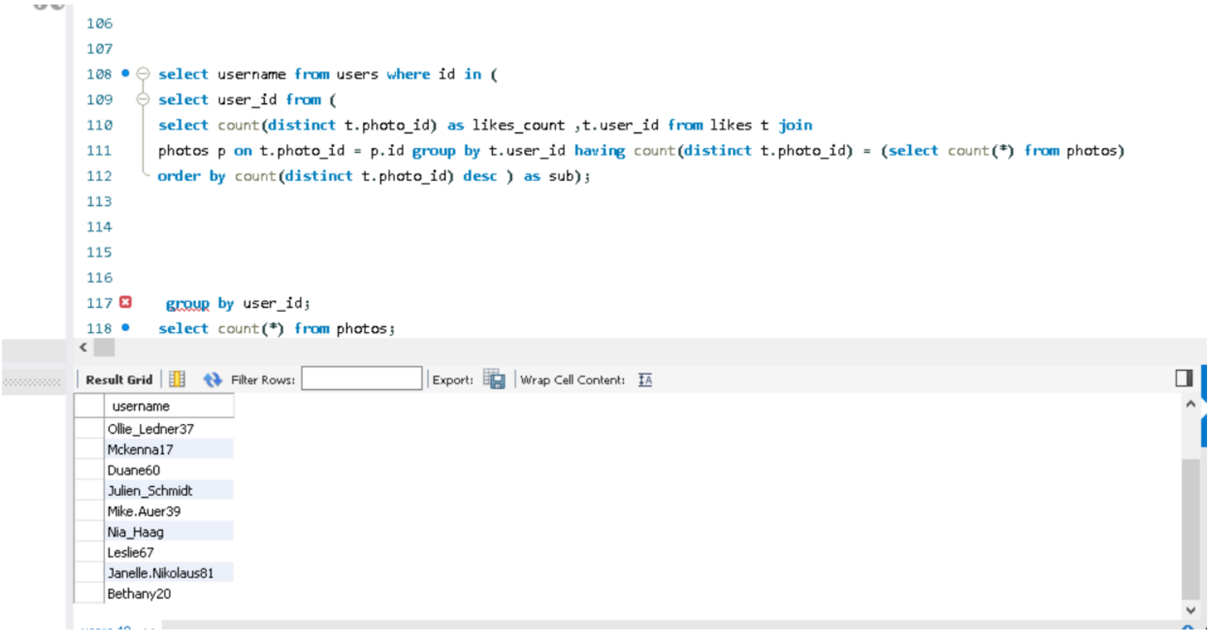
```
select round(avg(photo_id)) as avg_post, user_id from likes group by user_id ;

SELECT
    (SELECT COUNT(*) FROM photos) / (SELECT COUNT(*) FROM users) AS average_posts_per_user;
```

average_posts_per_user
2.5700

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

First i found the user id from likes tables i.e like count of every user and i found the total number of photos and matched the user's total likes count with total photo if matched it is a bot i.e that user liked every photo



```
select username from users where id in (
select user_id from (
select count(distinct t.photo_id) as likes_count ,t.user_id from likes t join
photos p on t.photo_id = p.id group by t.user_id having count(distinct t.photo_id) = (select count(*) from photos)
order by count(distinct t.photo_id) desc ) as sub);




 group by user_id;
select count(*) from photos;
```

username
Ollie_Ledner37
Mckenna17
Duane60
Julien_Schmidt
Mike.Auer39
Nia_Haag
Leslie67
Janelle.Nikolaus81
Bethany20