

Instagram user analytics



Agenda

- Project Description
- Approach
- Tech-Stack Used
- Insights
- Result



PROJECT DESCRIPTION

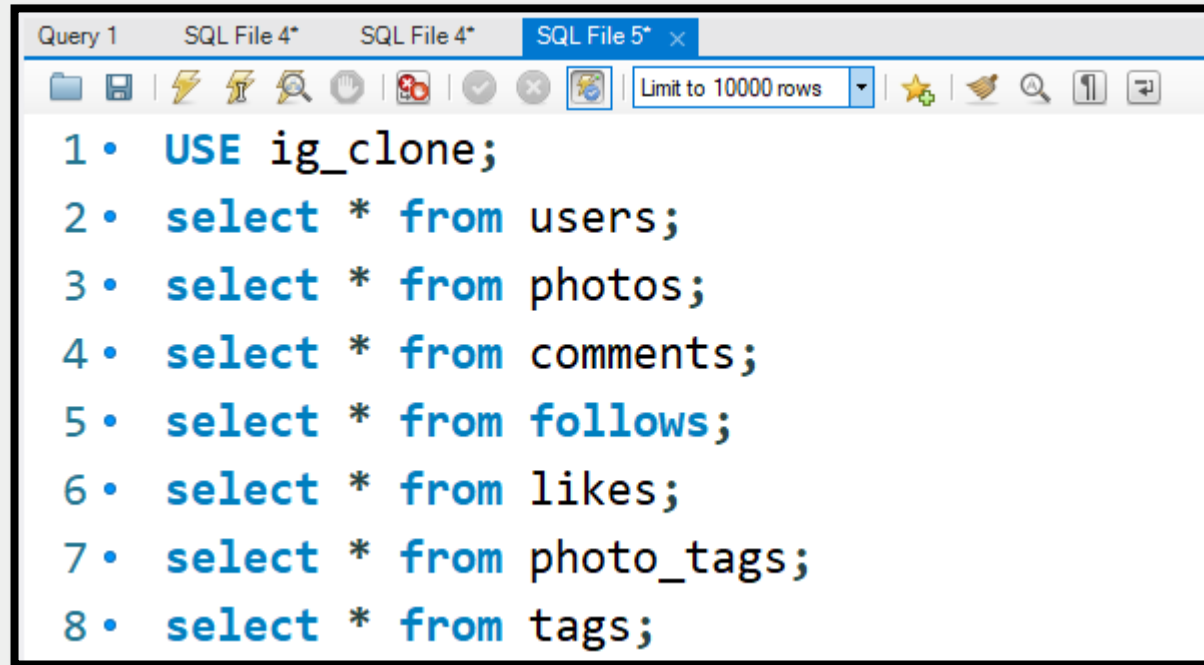
- This project focuses on analyzing user interactions and engagement with the Instagram platform to provide valuable insights to the marketing and investor teams. By leveraging data stored in the database, the goal is to uncover patterns, trends, and anomalies that can inform business decisions, improve user experience, and drive platform growth.
- **Purpose**
 1. **Marketing Team:** To identify loyal users, engage inactive users, and determine optimal strategies for campaigns, contests, and partnerships.
 2. **Investor Metrics:** To evaluate user engagement, detect potential fake accounts, and assess overall platform health.



APPROACH

1. Set Up the Database:

1. Load the provided database in MySQL Workbench and check the tables to understand the data by using the following queries.

A screenshot of the MySQL Workbench SQL editor interface. The window has a title bar with tabs for 'Query 1', 'SQL File 4*', 'SQL File 4*', and 'SQL File 5* x'. Below the title bar is a toolbar with various icons for file operations, execution, and viewing. A dropdown menu shows 'Limit to 10000 rows'. The main text area contains eight numbered SQL queries:

```
1 • USE ig_clone;  
2 • select * from users;  
3 • select * from photos;  
4 • select * from comments;  
5 • select * from follows;  
6 • select * from likes;  
7 • select * from photo_tags;  
8 • select * from tags;
```



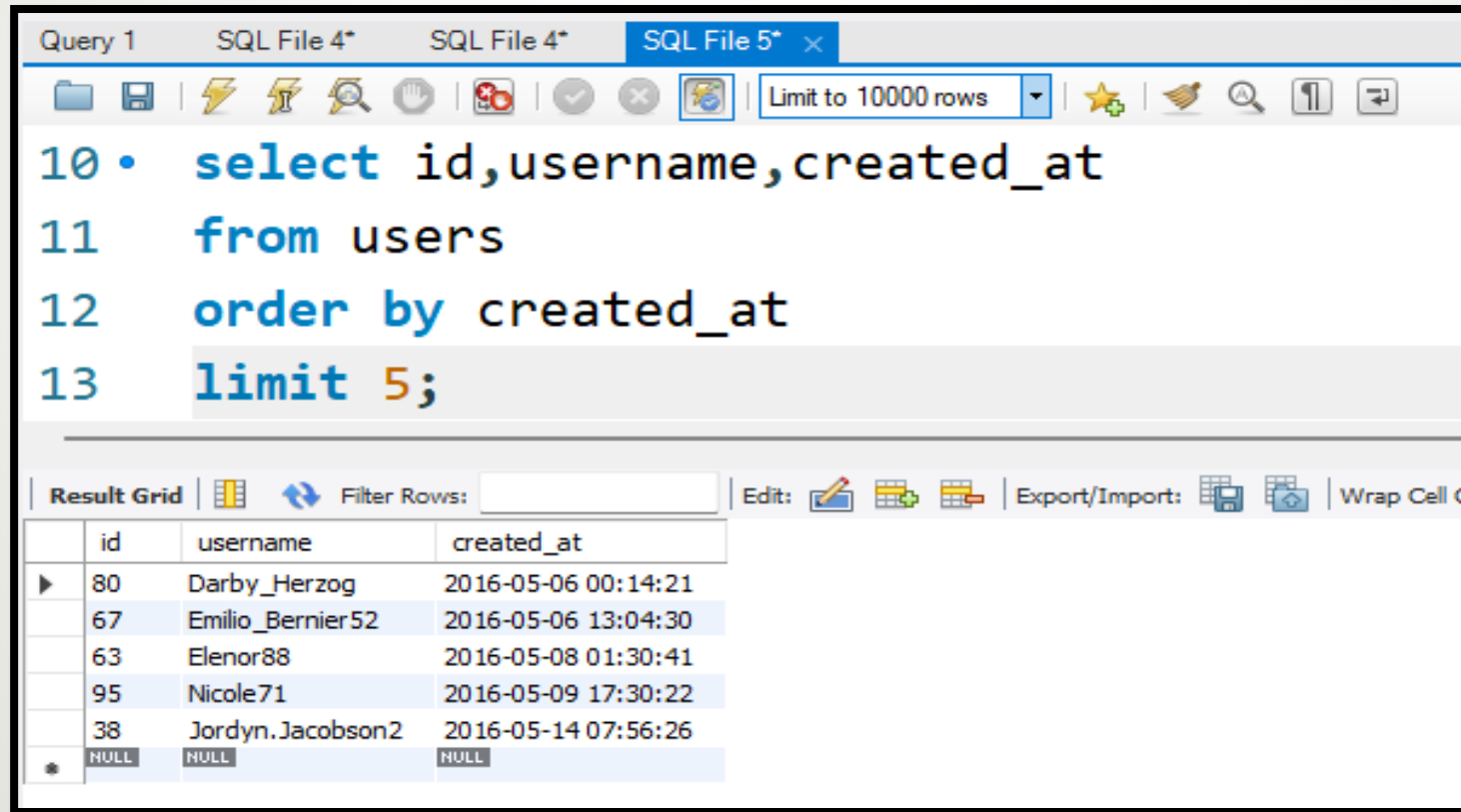
1. Write SQL Queries for the given Questions:

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.

ANS:



The screenshot shows a SQL IDE interface with a query editor and a result grid. The query editor contains the following SQL code:

```
10 • select id,username,created_at
11   from users
12   order by created_at
13   limit 5;
```

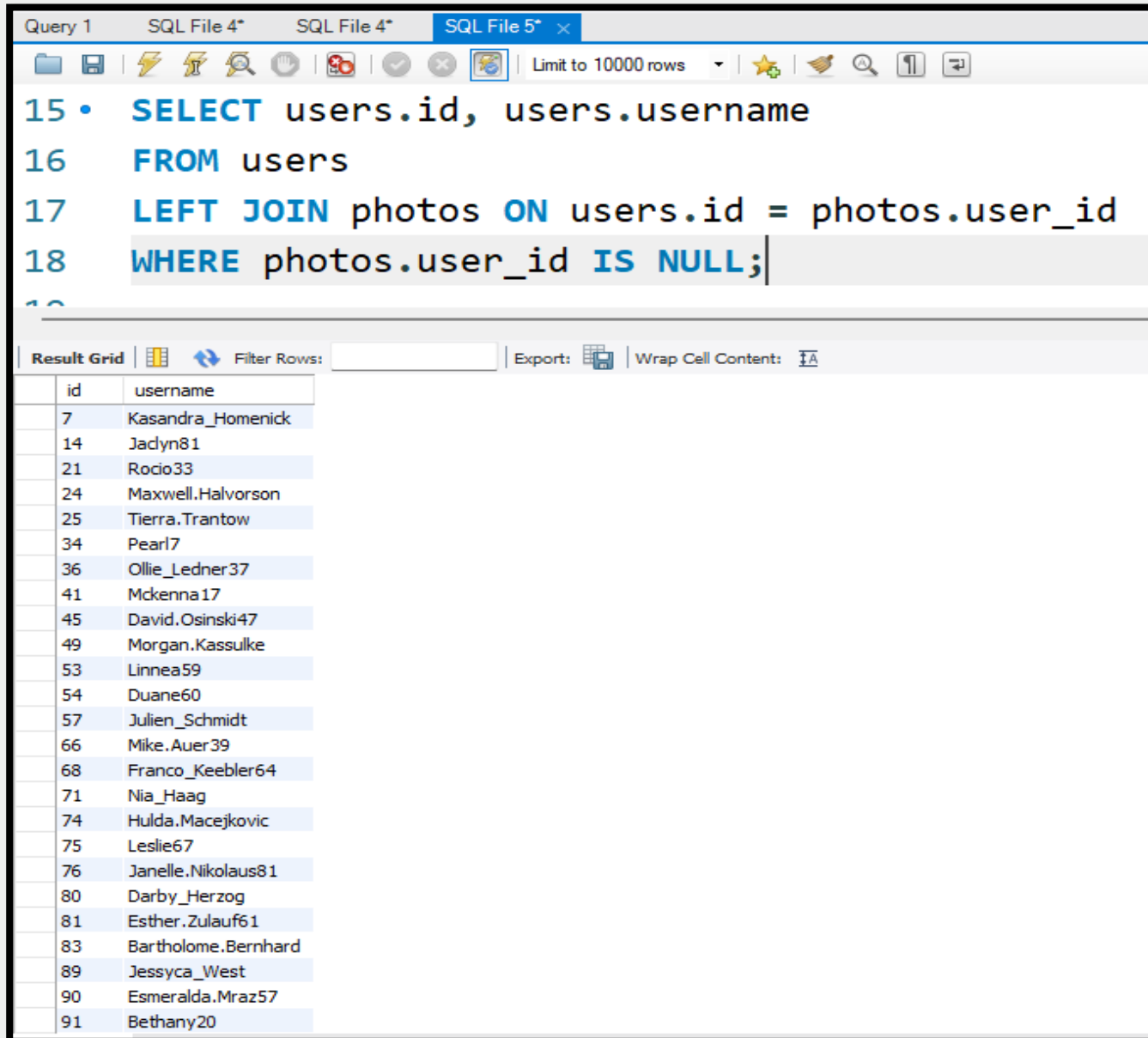
The result grid displays the following data:

	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



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.



The screenshot shows a SQL IDE interface with a query editor and a result grid. The query editor contains the following SQL code:

```
15 • SELECT users.id, users.username
16 FROM users
17 LEFT JOIN photos ON users.id = photos.user_id
18 WHERE photos.user_id IS NULL;
```

The result grid displays the following data:

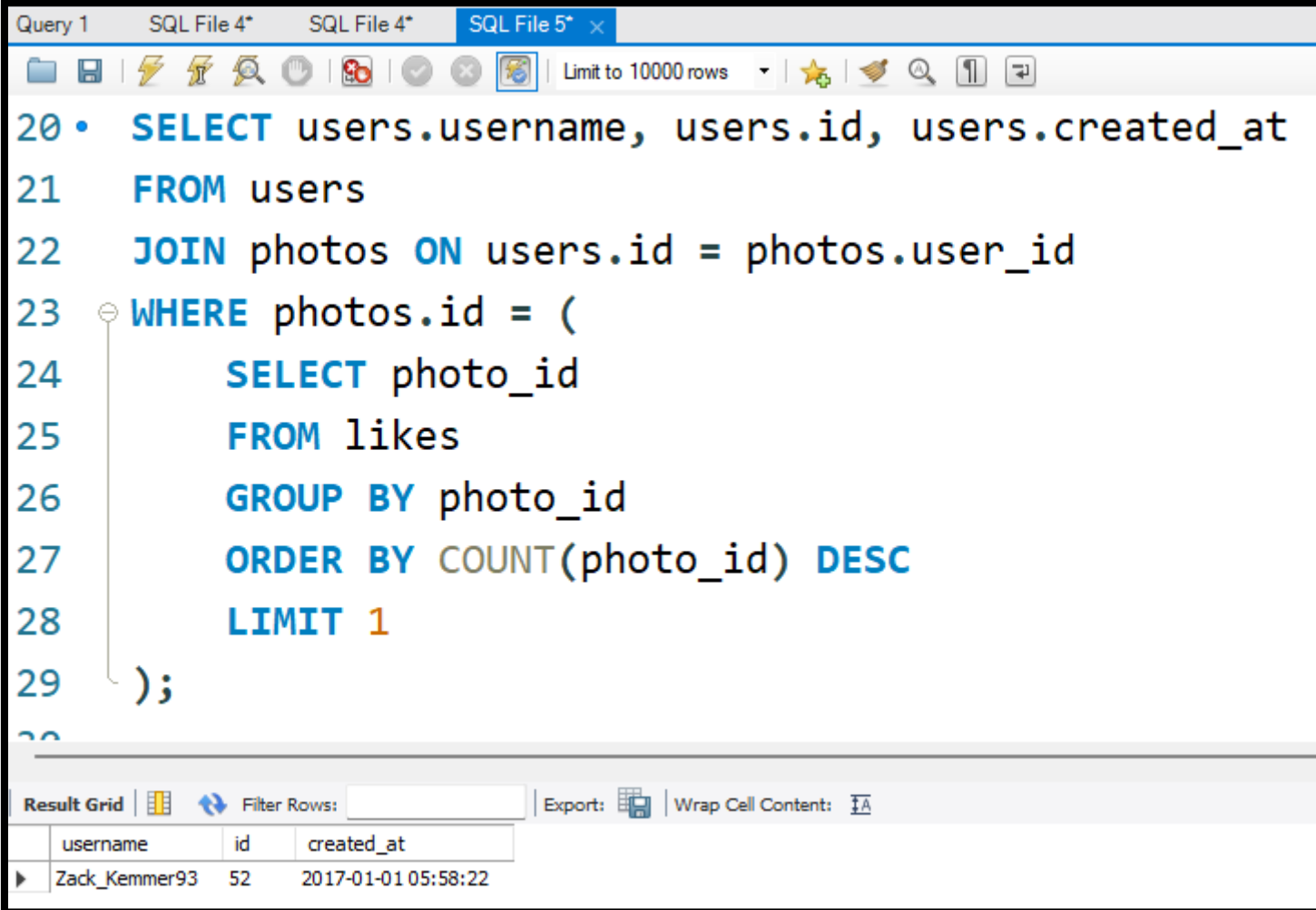
	id	username
7		Kassandra_Homenick
14		Jadyn81
21		Rocio33
24		Maxwell.Halvorson
25		Tierra.Trantow
34		Pearl7
36		Ollie_Ledner37
41		Mckenna17
45		David.Osinski47
49		Morgan.Kassulke
53		Linnea59
54		Duane60
57		Julien_Schmidt
66		Mike.Auer39
68		Franco_Keebler64
71		Nia_Haag
74		Hulda.Macejkovic
75		Leslie67
76		Janelle.Nikolaus81
80		Darby_Herzog
81		Esther.Zulauf61
83		Bartholome.Bernhard
89		Jessyca_West
90		Esmeralda.Mraz57
91		Bethany20



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

ANS:



```
Query 1  SQL File 4*  SQL File 4*  SQL File 5* x
Limit to 10000 rows
20 • SELECT users.username, users.id, users.created_at
21 FROM users
22 JOIN photos ON users.id = photos.user_id
23 WHERE photos.id = (
24     SELECT photo_id
25     FROM likes
26     GROUP BY photo_id
27     ORDER BY COUNT(photo_id) DESC
28     LIMIT 1
29 );
```

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

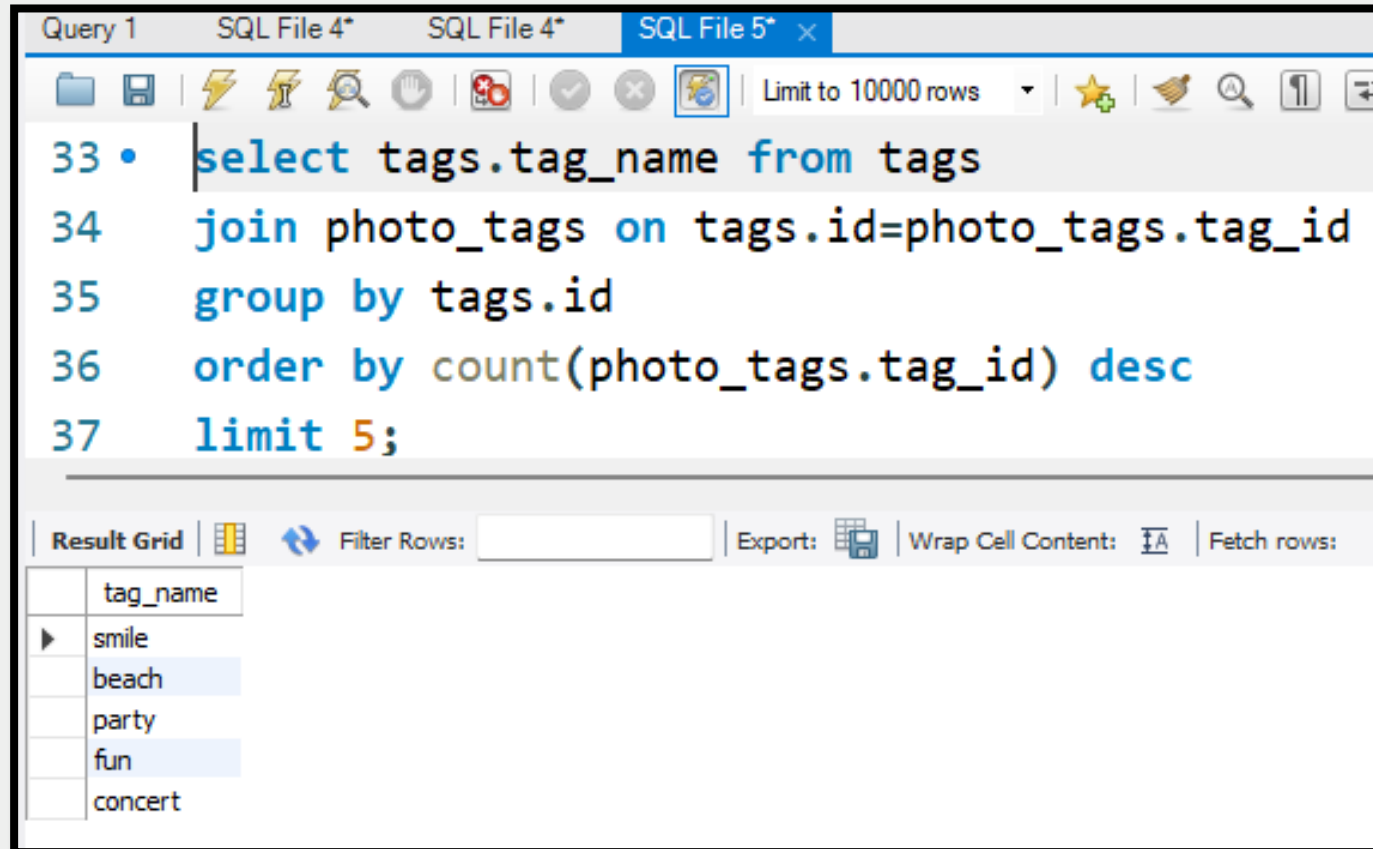
username	id	created_at
Zack_Kemmer93	52	2017-01-01 05:58:22



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.

ANS:



The screenshot shows a SQL IDE window with a query editor and a result grid. The query editor contains the following SQL code:

```
33 • select tags.tag_name from tags
34   join photo_tags on tags.id=photo_tags.tag_id
35   group by tags.id
36   order by count(photo_tags.tag_id) desc
37   limit 5;
```

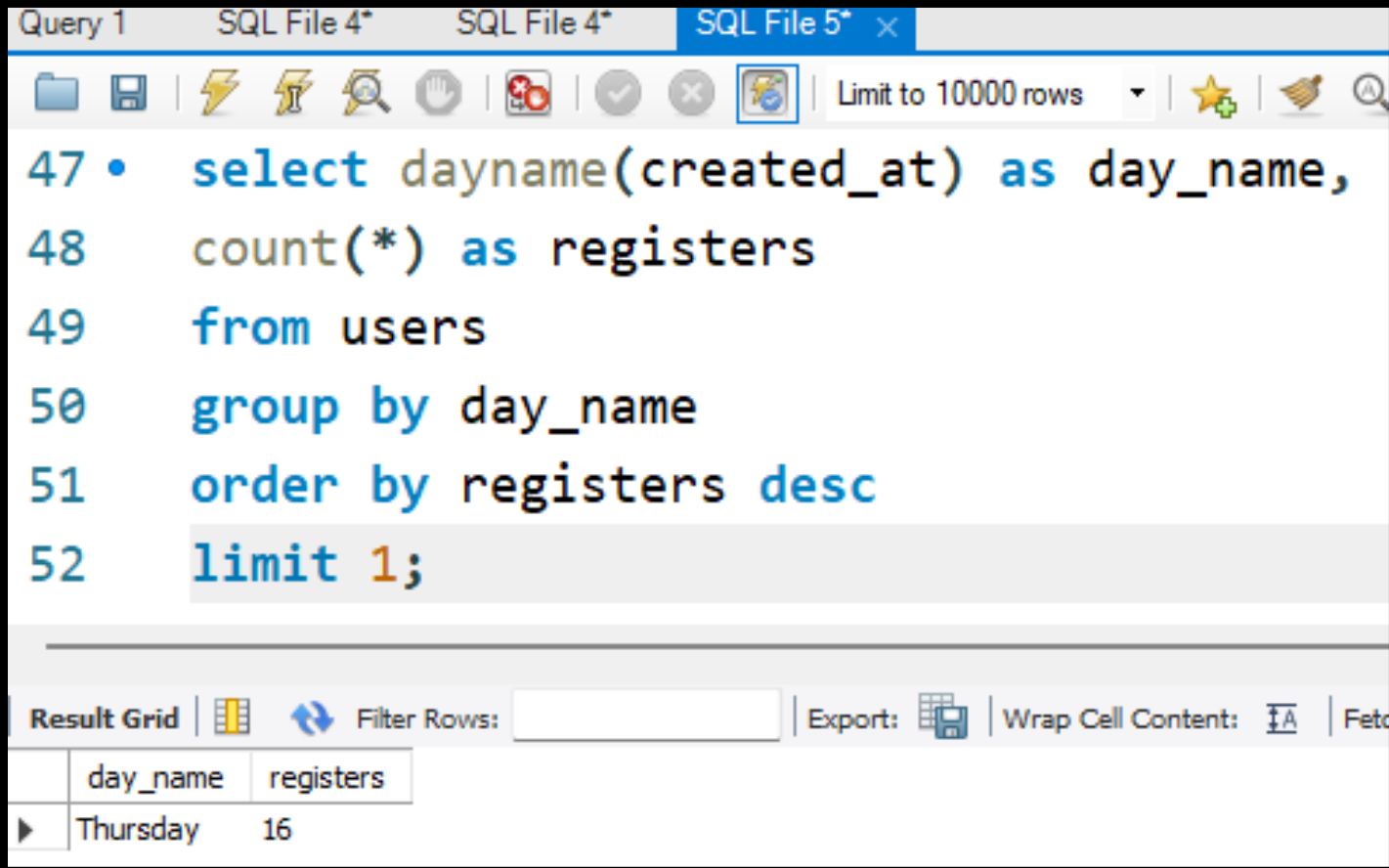
The result grid displays the following data:

	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.

ANS :



The screenshot shows a SQL IDE window with multiple tabs. The active tab is 'SQL File 5*'. The query editor contains the following SQL code:

```
47 • select dayname(created_at) as day_name,  
48      count(*) as registers  
49      from users  
50      group by day_name  
51      order by registers desc  
52      limit 1;
```

Below the query editor, the 'Result Grid' is visible. It shows a table with two columns: 'day_name' and 'registers'. The first row of data is 'Thursday' with a value of 16.

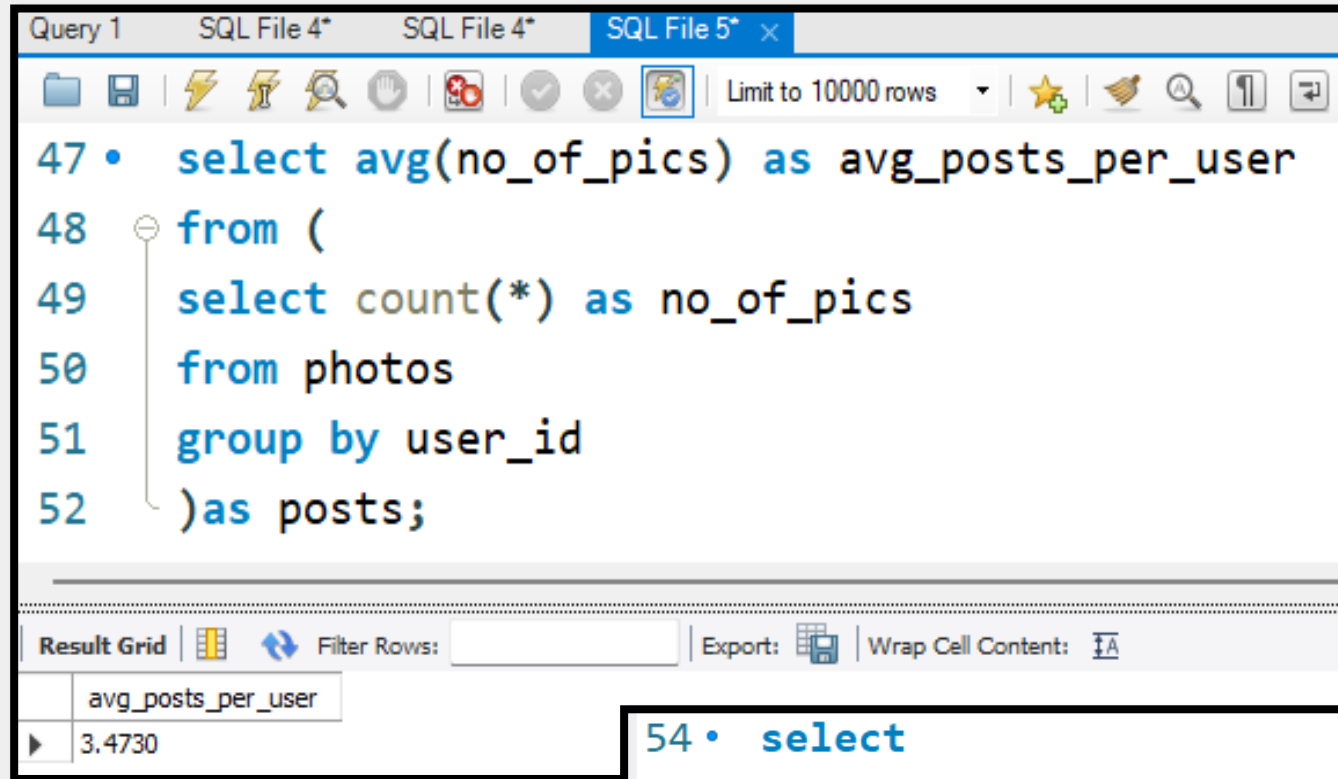
day_name	registers
Thursday	16



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.

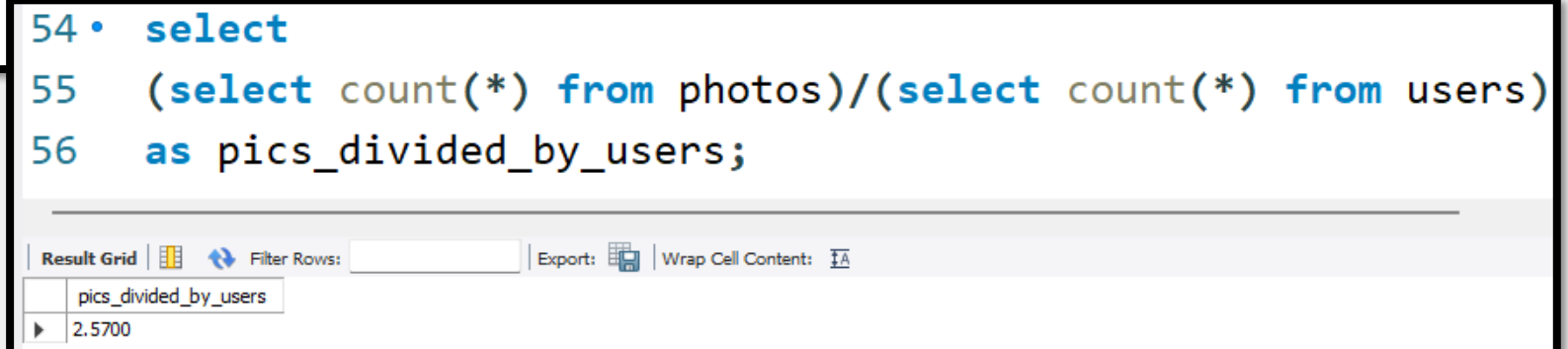


The screenshot shows a SQL IDE window with a query editor and a result grid. The query is as follows:

```
47 • select avg(no_of_pics) as avg_posts_per_user
48   from (
49     select count(*) as no_of_pics
50     from photos
51     group by user_id
52   )as posts;
```

The result grid shows the following data:

avg_posts_per_user
3.4730



The screenshot shows a SQL IDE window with a query editor and a result grid. The query is as follows:

```
54 • select
55   (select count(*) from photos)/(select count(*) from users)
56   as pics_divided_by_users;
```

The result grid shows the following data:

pics_divided_by_users
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.

```
58 • select user_id, count(*) as Tot_likes from likes
59   group by user_id
60   having Tot_likes= (select count(*) from photos);
```

Result Grid			Filter Rows:	Export:	Wrap Cell Content:
	user_id	Tot_likes			
▶	5	257			
	14	257			
	21	257			
	24	257			
	36	257			
	41	257			
	54	257			
	57	257			
	66	257			
	71	257			
	75	257			
	76	257			
	91	257			



TECH STACK USED

1.MySQL Workbench

1. Version: 8.0

2. Why I Used It:

MySQL Workbench made it super easy to manage the database and run SQL queries. Its simple interface and visual tools helped me analyze the data quickly and efficiently.

2.MySQL Server

1. Version: 8.0

2. Why I Used It:

MySQL Server is reliable and works smoothly with MySQL Workbench. It handled all the data without any issues, making the whole process hassle-free.

3.Operating System

1. Used: Windows 11

2. Why I Used It:

I worked on Windows 11 because it's easy to set up and works well with the tools I needed for the project.



INSIGHTS

While working on the project, I gained valuable insights about user interactions and engagement on the platform. Here are the key findings:

1.Loyal Users

1. The five oldest users were identified, showing who has been using the platform the longest. These users could be rewarded to boost loyalty.

2.Inactive Users

1. A significant number of users have never posted a single photo. This highlights an opportunity to engage these users through targeted campaigns.

3.Contest Winner

1. The photo with the highest likes was identified, along with the user who posted it. This user could be recognized to promote engagement.

4.Popular Hashtags

1. The top five hashtags were determined, providing insights for brands and users to increase their reach on the platform.



5.Best Day for Ad Campaigns

Most users registered on a specific day of the week, suggesting the best time to launch ad campaigns for maximum visibility.

6.User Engagement Metrics

The average number of posts per user was calculated, helping to understand overall engagement levels on the platform.

7.Bots and Fake Accounts

A few suspicious users who liked every photo were flagged as potential bots, indicating the need for better monitoring of user activity.



RESULT

Through this project, I was able to achieve the following:

1.Valuable Insights for Marketing and Product Teams

- Identified the most loyal users who can be rewarded to build stronger relationships.
- Found inactive users, providing an opportunity to re-engage them with targeted campaigns.

2.Investor-Friendly Metrics

- Delivered key engagement metrics like average posts per user and flagged potential bot accounts, which can help improve platform credibility and user experience.

3.Clear Analytical Process

- Gained hands-on experience writing SQL queries, analyzing data, and presenting results in an easy-to-understand manner.
- Learned how to approach data analysis projects systematically and derive meaningful conclusions.



Impact of the Analysis

The insights derived from this project can significantly benefit the platform. Marketing teams can launch better campaigns, product teams can focus on enhancing user activity, and investors can rely on meaningful engagement metrics to assess platform health. Personally, this project enhanced my analytical and SQL skills, providing me with a deeper understanding of data-driven decision-making.



