INSTAGRAM USER ANALYTICS

(SQL FUNDAMENTALS)

PROJECT DESCRIPTION

In this project I'll do a work of data analyst. We'll analyze with user interactions and engagement with the Instagram app to provide valuable insights that can help business grow.

Also we'll track how users engage with a digital product such as a software or a mobile application or a software application.

Those things will help in many ways i.e the marketing team might use these insights to launch a new campaign, the product team might use them to decide on new features to build, and the development team might use them to improve the overall experience.

In this project I'll be using MySQL to analyze **Instagram data** and my insights will help the product manager and rest of the team to make decisions about the future features on the Instagram app.

APPROACH

Firstly, I'll take approach to upload the dataset in MySQL then create and insert the values in the database by using the DDL (Data Definition Language), DML (Data Manipulation Language) by using MySQL workbench.

TECH – STACK USED

The MySQL Workbench Community and MySQL and MySQL community serverand MySQLWorkbench release for versions 8.0 through 8.0.34.

INSIGHTS

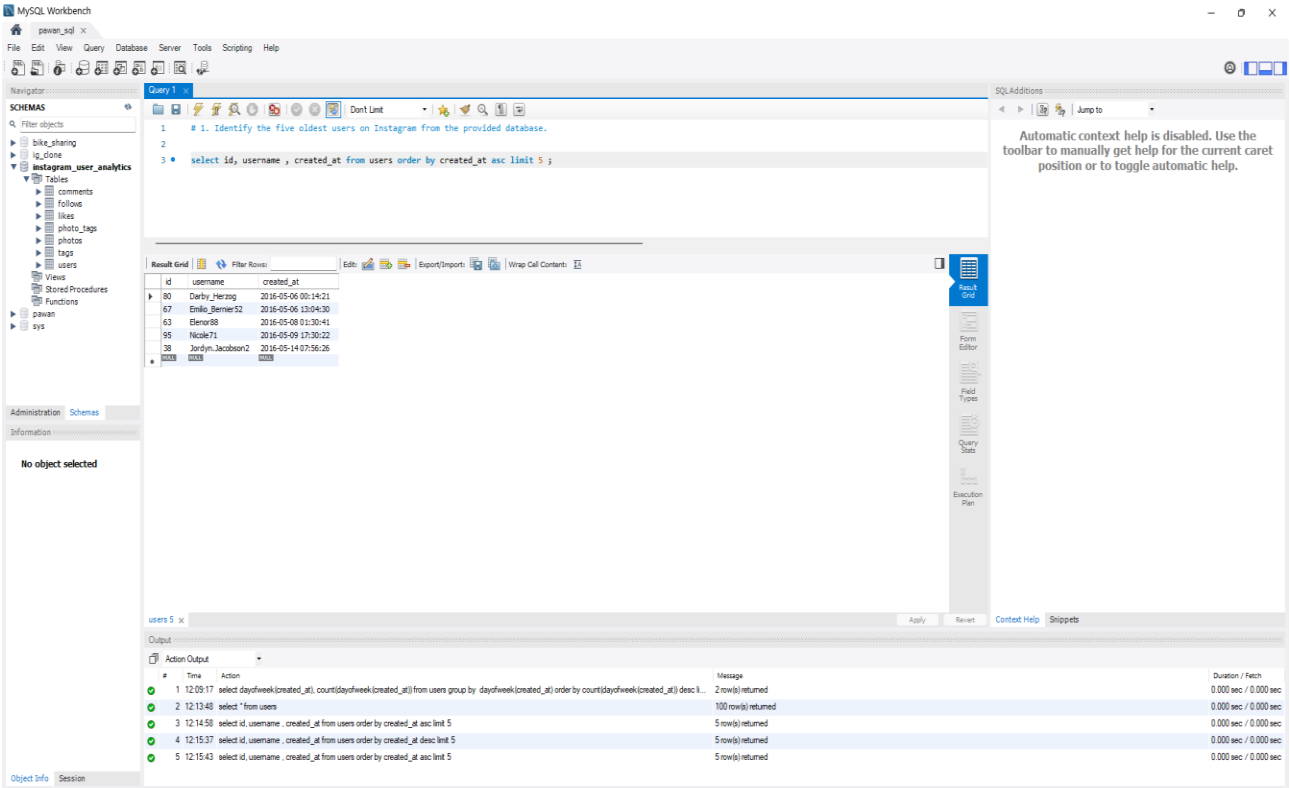
A . Marketing Analysis :

The marketing team might use these insights to launch a new campaign.

1. Identify the five oldest users on Instagram from the provided Database.

In this question we have to analyze the five oldest users on Instagram from the database because the marketing team wants to reward the most loyal users i.e those who have been using the platform for the longest time.
So firstly, we select the **id , username, created_at** from users table as this will help to extract the required output from the database after this we'll arrange the data in descending order by using **order by desc** approach and we'll set the **limit 5** because we need only the top 5 oldest Instagram accounts from database.

Here's our SQL output :



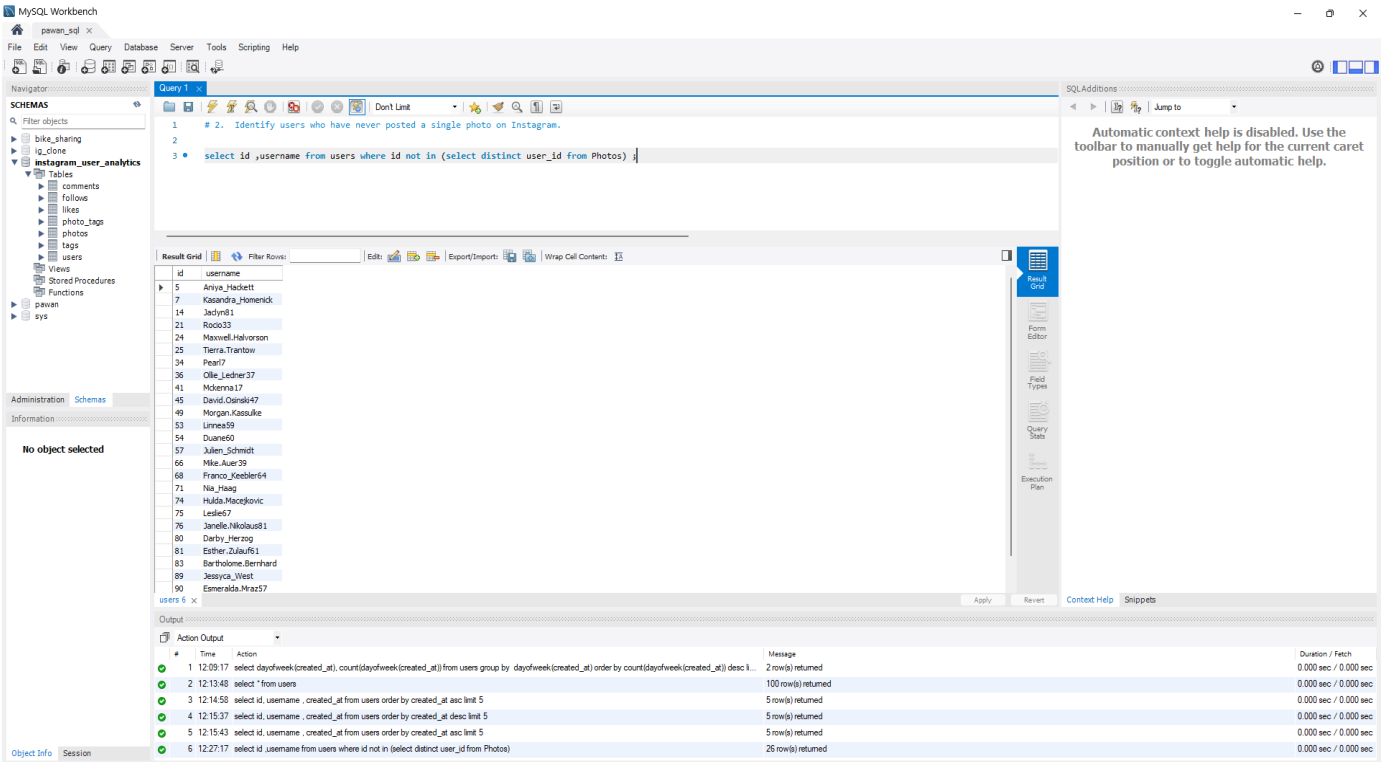
2. Identify the users who have never posted a single photo on Instagram.

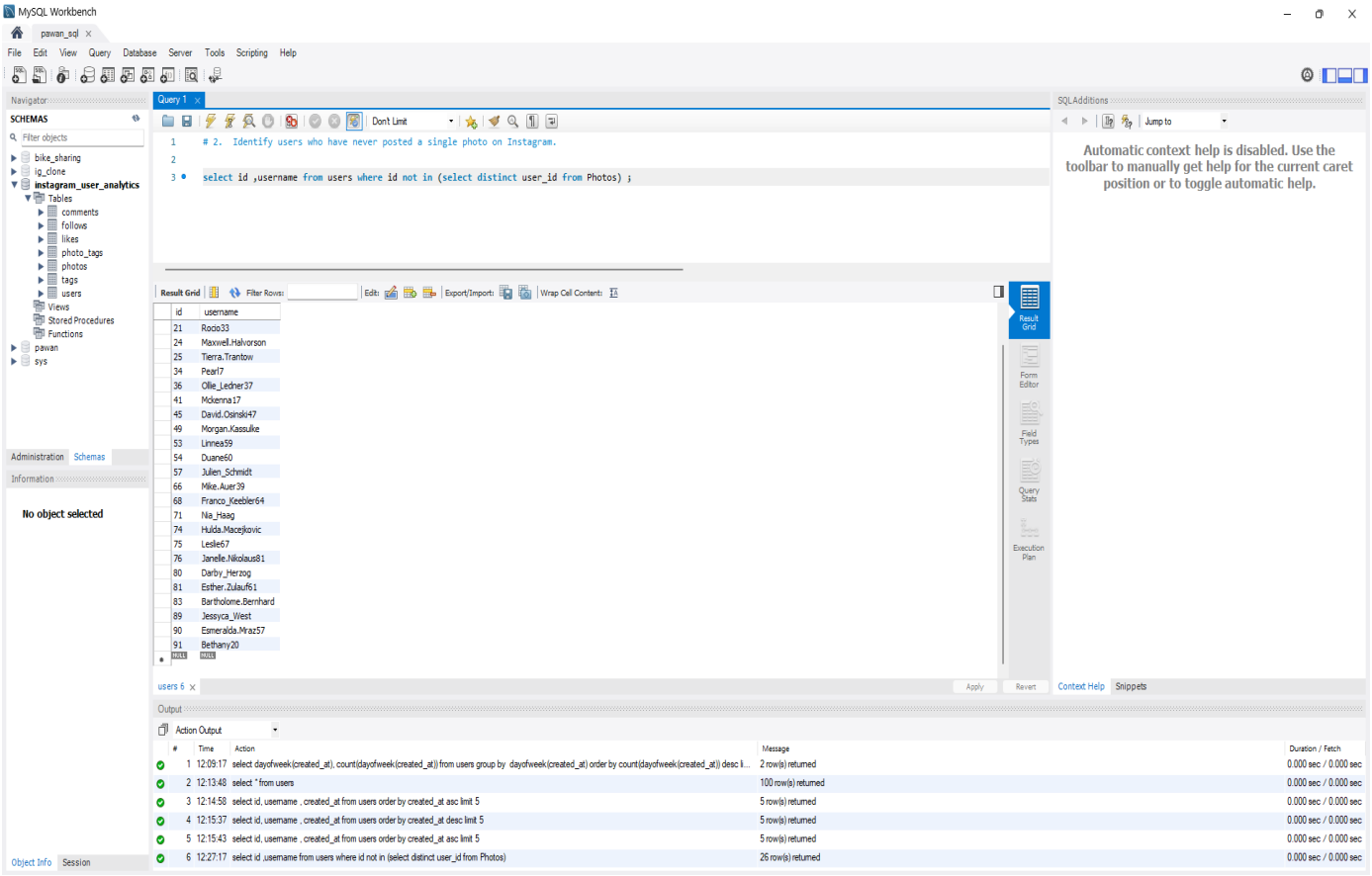
In this question, we want those users who have never posted a single photo on Instagram so the team wants to encourage inactive users to start posting by sending them promotional emails. So firstly we **select distinct ids from table Photos** which we'll use to compare id's from **users** table.

Now, we **select id, username from users** where **id** not in the selected ids from Photos table**.**

We did this because there won't be any photos posted by a person whose id's not active.

Here's our output :





3. Determine the winner of the contest and provide their details to the team.

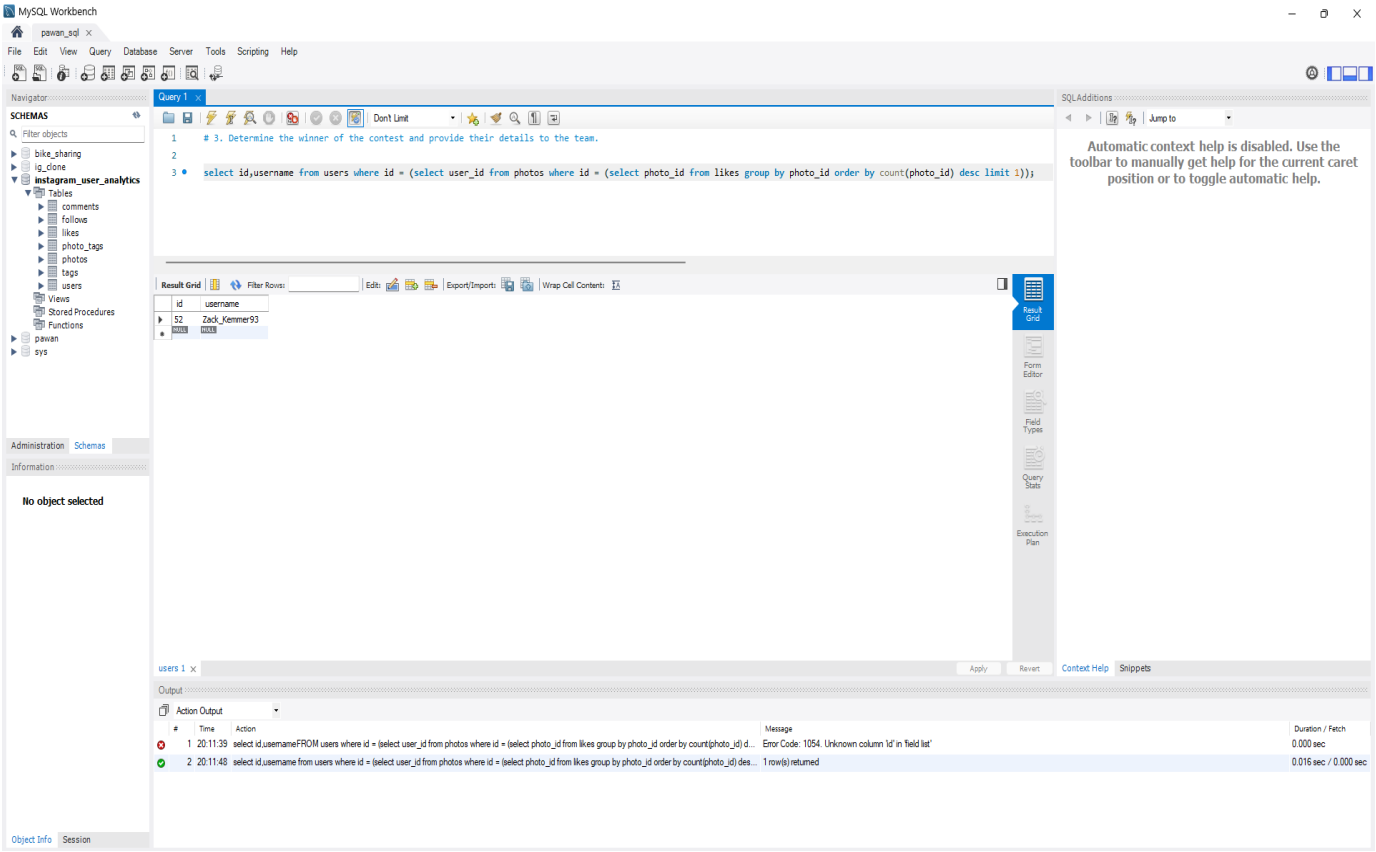
In this question, we have to analyze the winner of the contest who get the most likes in the single photo because the team has organized a contest where the user with the most likes on a single photo wins.

So firstly, we select **the photo_id from likes** and **group by photo_id** then **order by count of photo_id** and arranging them in descending order by using **desc** syntax and we set the **limit to 1**.

We're taking this approach to get output of the user photo which gets most likes in a single photo.

Now we **select id, username from users** and use where function with **id** and we'll get the required output.

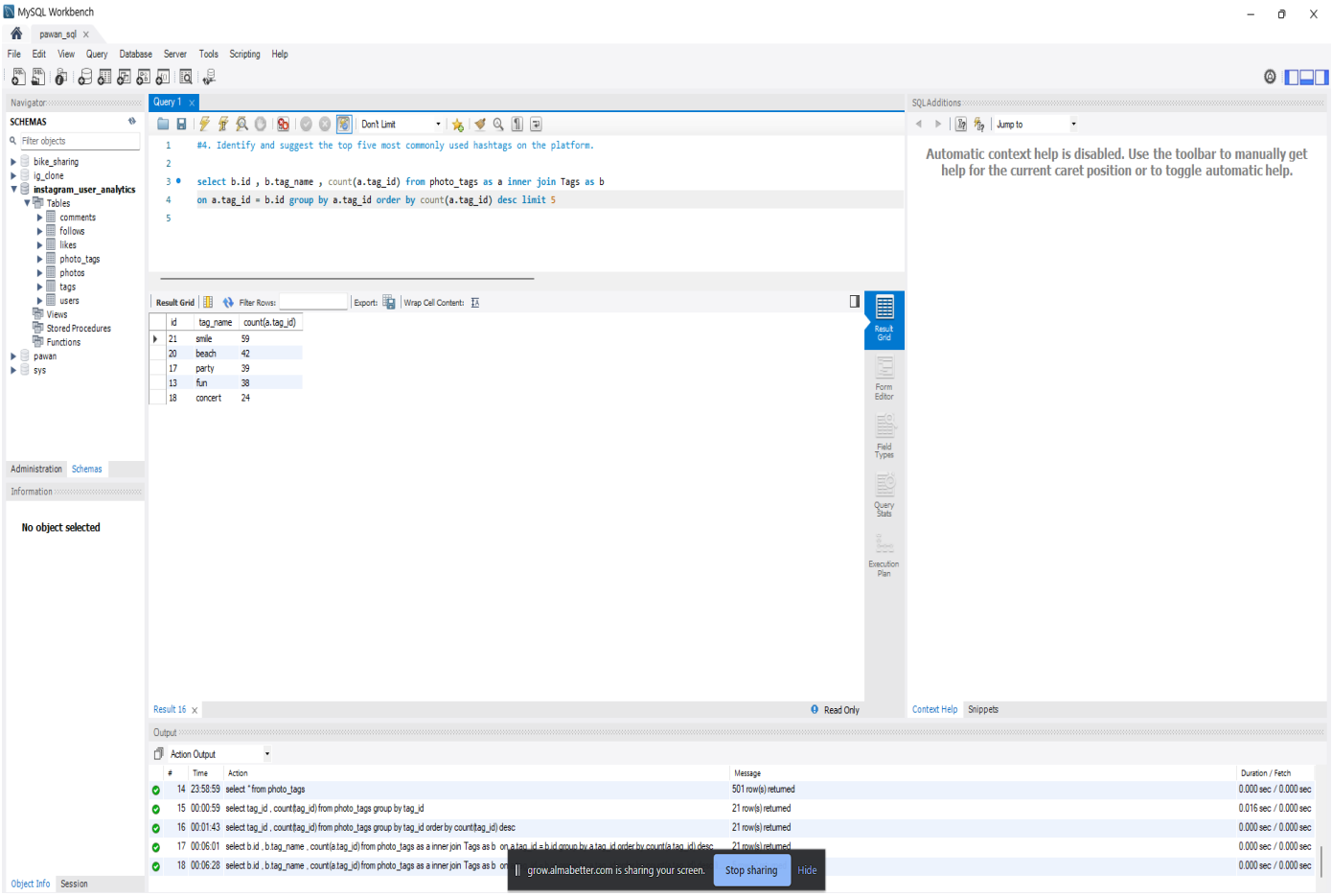
Here's our output :



4. Identify and suggest the top five most commonly used hashtags on the platform.

In this question, we analyze the top five most commonly used hashtags because the partner brand wants to know the most popular hashtags to use in their posts to reach the most people.

For this we will be using **inner join** to join the tables **photo_tags** and tags on the condition of equitable **ids** also grouping by **tag_id** and ordering by **tag_id.** Now, by arranging the output generated in descending order by using **desc** and limiting the record count to 5, we get the desired output.

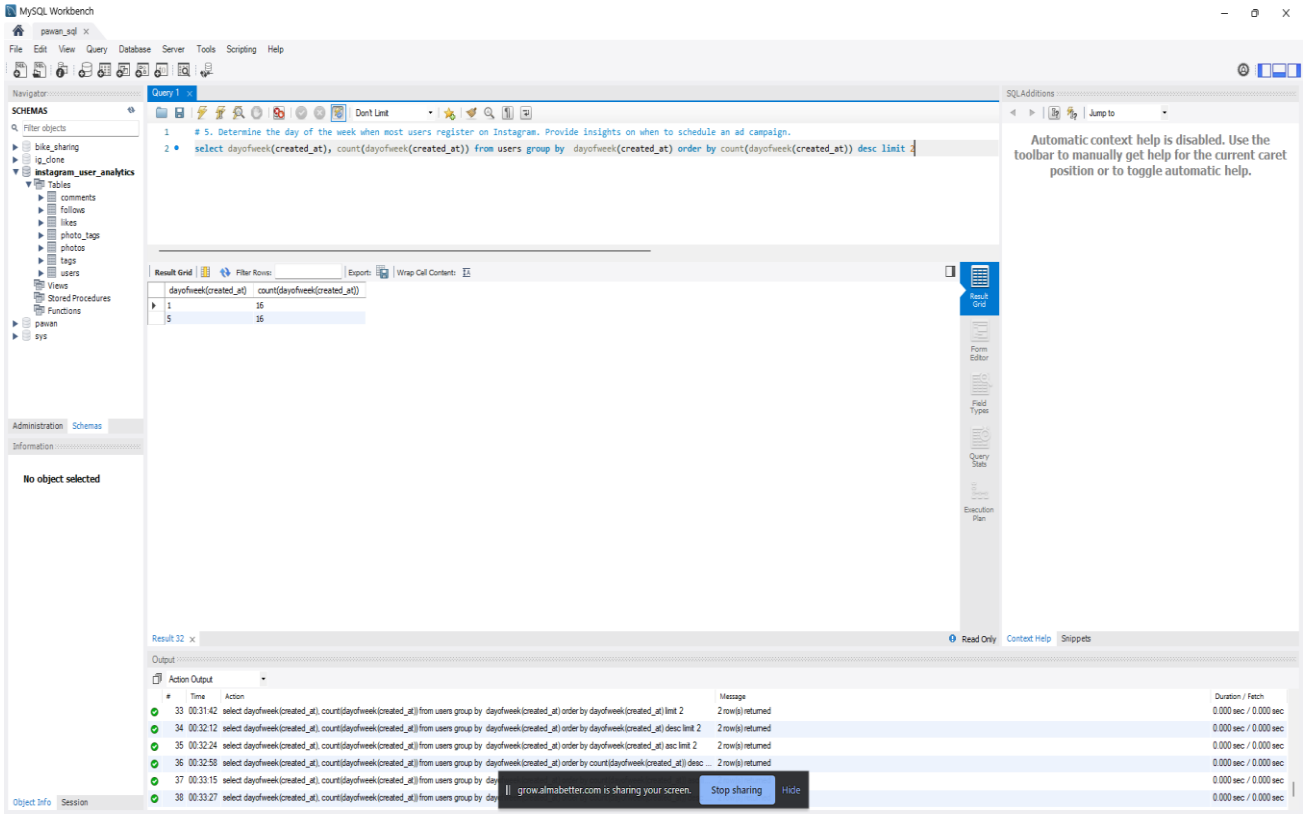


5. Determine the day of the week when the most users register on Instagram. Provide insights on when to schedule on ad campaign.

In this question, we analyze the day when the most users register on Instagram because the team want to know the best day of the week to launch ads.

Firstly, we **select** the **dayofweek of created_at** and **count** this insights from **users** and after using this we'll apply **group by** function with **dayofweek of created_at** and after this we arrange in descending order by using **desc** and set **limit 2**.

Here's our output :



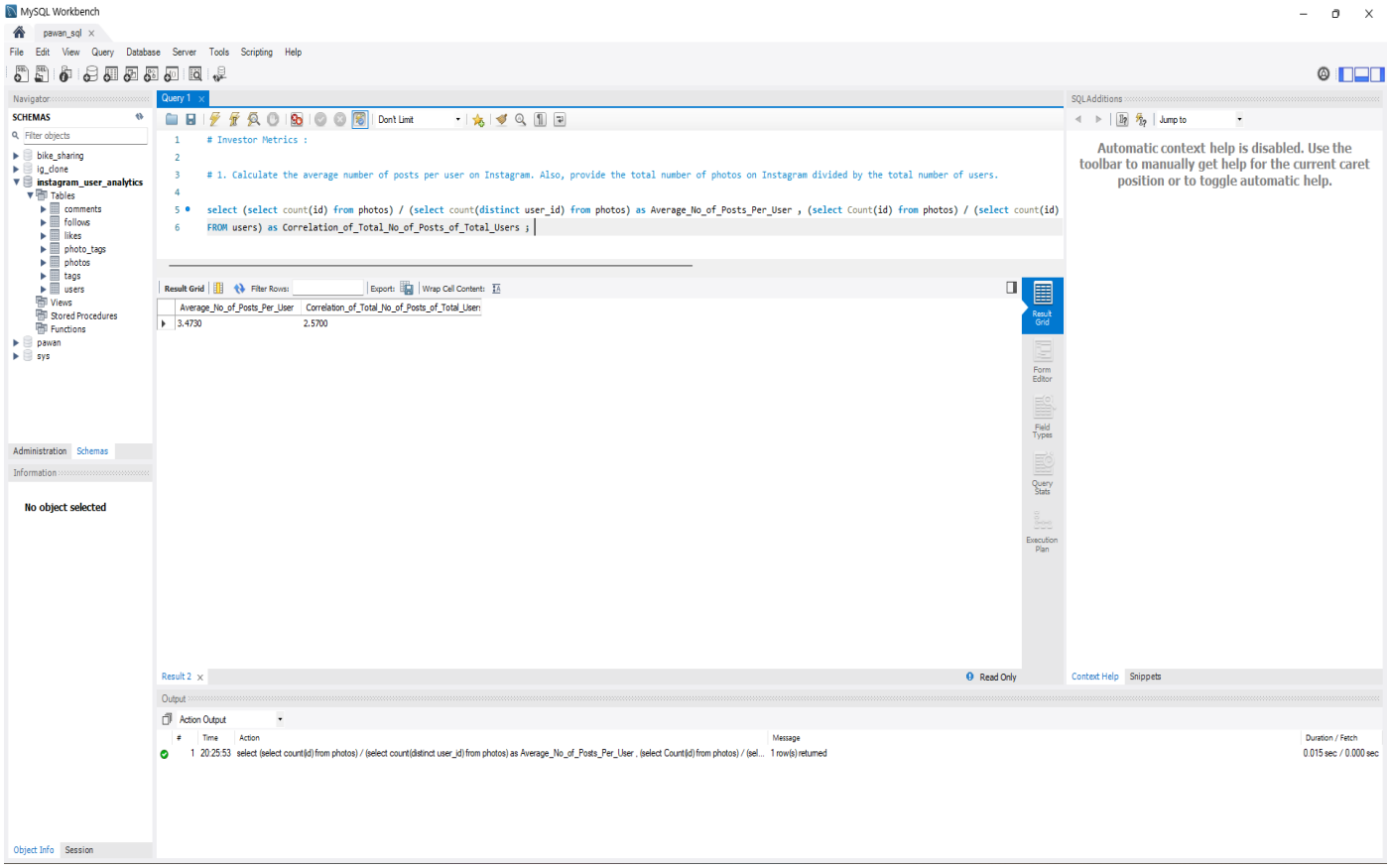
B. Investor Metrics :

1. **Calculate the average number of posts per user on Instagram. Also, provide the total number of photos on Instagram divided by the total numbers of users.**

In this question we have to analyze the average number of posts per users on Instagram and the correlation between the total numbers of photos and the total numbers of users because Investor wants to know if users are still active and posting on Instagram or if they are making fewer posts.

Firstly, **select** the count of **id** from **photos** and divide by the count of **user_id** from **photos** to get **average number of posts of users** and after this we select the count of **id** from **photos** and divide by count of **id** from **users** so after doing this we get the **correlation between of total number of posts of total users.**

Here's our output :

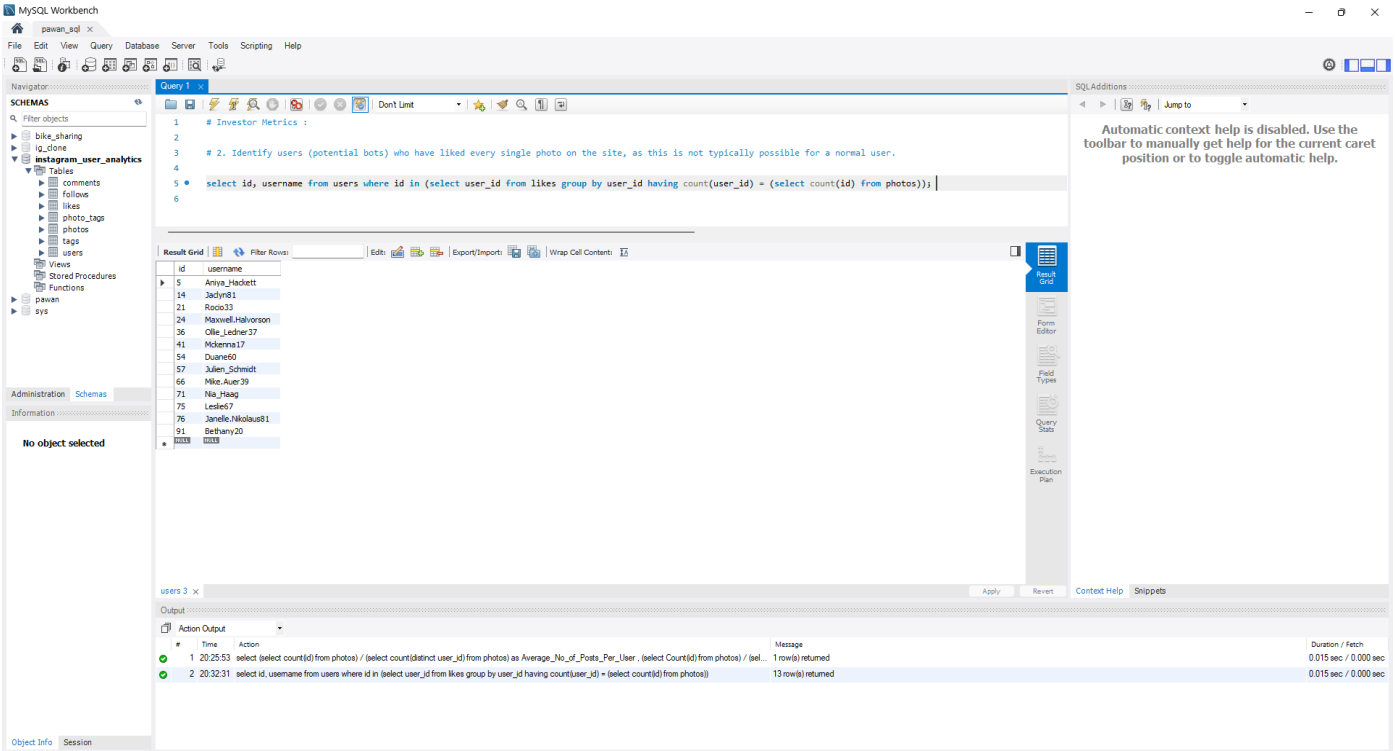


2. Identify users (potential bots) who have liked every single photo on the site , as this is not typically possible for a normal user.

In this question, we analyze the users who have liked the every single photo which is not typically possible for a normal user because the Investors want to know if the platform is crowded with fake and dummy accounts.

Firstly, we need to figure out the **ids** for which the **count of ids** from **likes** and **Photos** table are equal and also for doing this we **group by user id.** Finally we select **id** and **username** from **users** table to get the required details.

Here’s our output :



RESULTS

- After doing this project I learnt the fundamentals and got some knowledge how the data analyst works using MySQL and how they are using database to extract the output.
- I also learnt from this database and after this project how users engage with software application or a mobile application and after doing analysis of this data, I also got knowledge how the marketing team, software team and product team works together and derive the insights for business and how the data can be used to increase and optimize the application and engagement with users and audience.