Instagram User Analytics

Project Description

For this project, I analyzed Instagram user data using SQL and MySQL Workbench in order to respond to inquiries from the management team. My observations will assist the team's product manager and other members in making defensible choices on the Instagram app's future.
This project aims to leverage my SQL expertise to get insightful conclusions from the data. My research could have an impact on how one of the most widely used social media platforms in the world develops in the future.

Approach

The approach towards this project is to use SQL queries to analyze the database provided.

Tech-Stack Used

MySQL Workbench 8.0 CE is the program that I used to do this project. It is employed for writing SQL queries and building databases. Additionally, I make Word file using Microsoft Word and store SQL queries and snapshots of output in the same.

Insights

Below are the insights that I have brought into light for the management team for their clear understanding about the analytics for the instagram users data.

Top 5 Oldest Users of Instagram
Users who never posted photos on Instagram
Most liked photo on Instagram
Top 5 most commonly used has-tags on Instagram
Total number of users on Instagram
Total number of photos on Instagram
Average number of photos per user
Bots and fake accounts on Instagram

SQL Tasks :

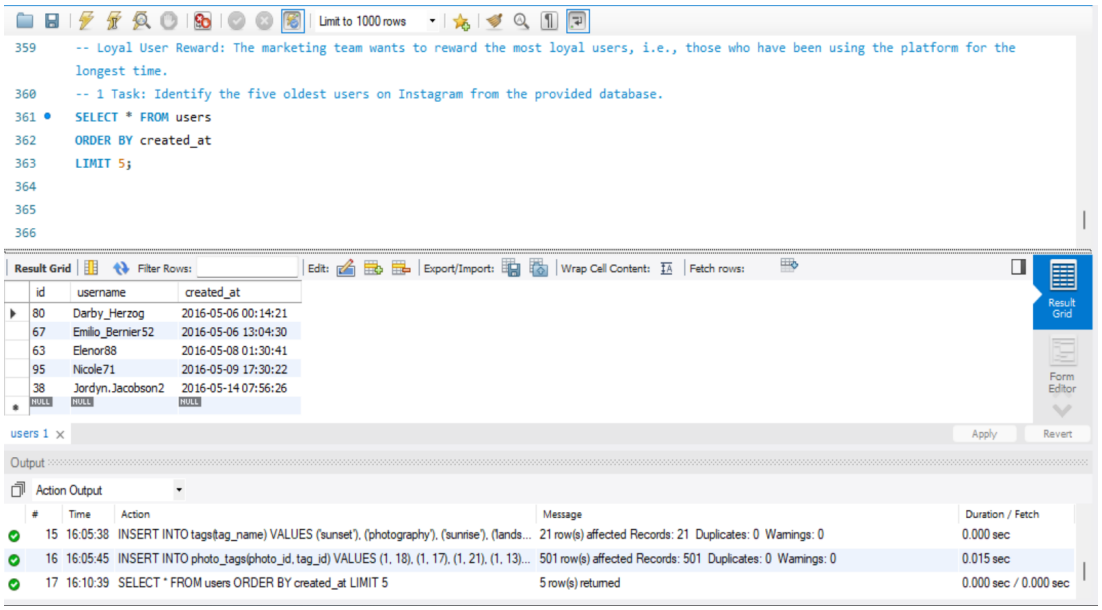
A) Marketing Analysis:

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

CODE:

```
SELECT * FROM users

ORDER BY created_at

LIMIT 5;
```

OUTPUT:

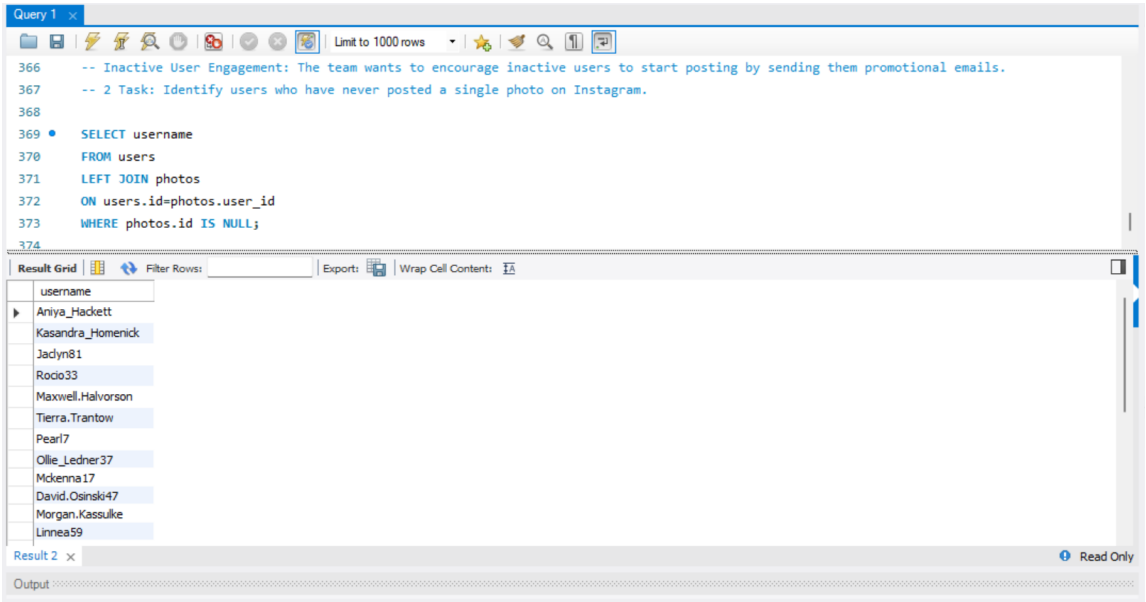


2. Inactive User Engagement: The team wants to encourage inactive users to start posting by sending them promotional emails.
Your Task: Identify users who have never posted a single photo on Instagram.

CODE:

```
SELECT username

FROM users

LEFT JOIN photos

ON users.id=photos.user_id

WHERE photos.id IS NULL;
```

OUTPUT:

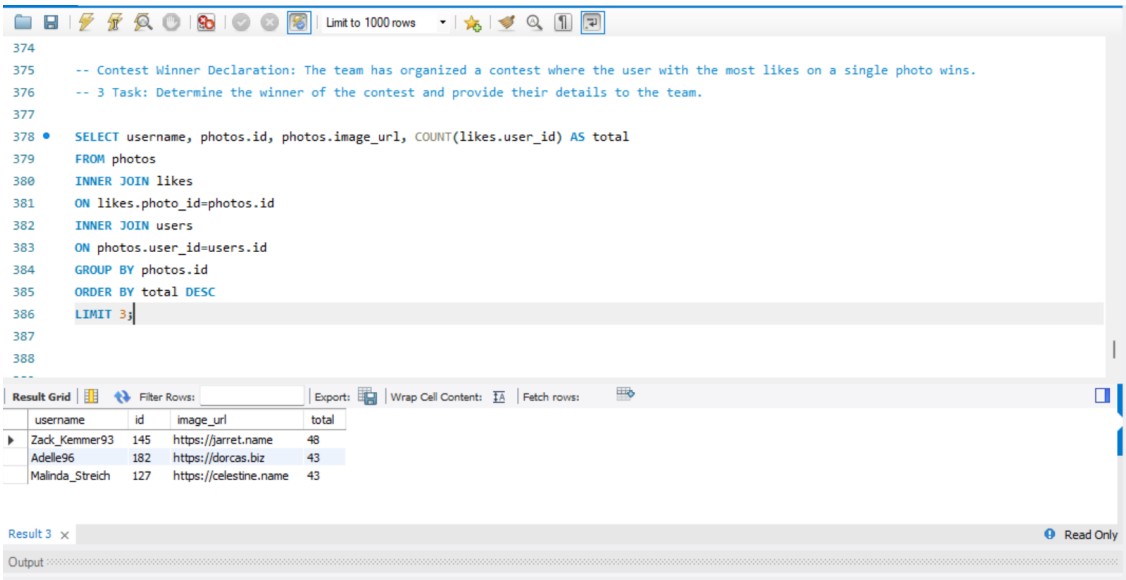


3. Contest Winner Declaration: The team has organized a contest where the user with the most likes on a single photo wins.
Your Task: Determine the winner of the contest and provide their details to the team.

CODE:

```
SELECT username, photos.id, photos.image_url, COUNT(likes.user_id) AS total

FROM photos

INNER JOIN likes

ON likes.photo_id=photos.id

INNER JOIN users

ON photos.user_id=users.id

GROUP BY photos.id

ORDER BY total DESC

LIMIT 3;
```

OUTPUT:

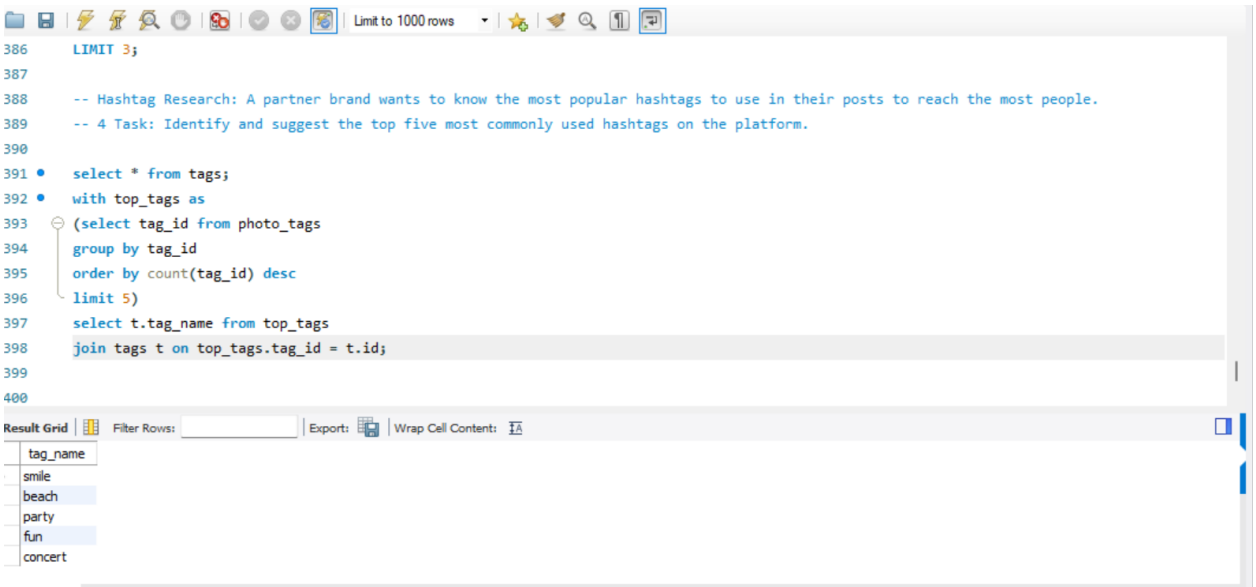


4. Hashtag Research: A partner brand wants to know the most popular hashtags to use in their posts to reach the most people.
Your Task: Identify and suggest the top five most commonly used hashtags on the platform.

CODE:

```
select * from tags;

with top_tags as

(select tag_id from photo_tags

group by tag_id

order by count(tag_id) desc

limit 5)

select t.tag_name from top_tags

join tags t on top_tags.tag_id = t.id;
```

OUTPUT:

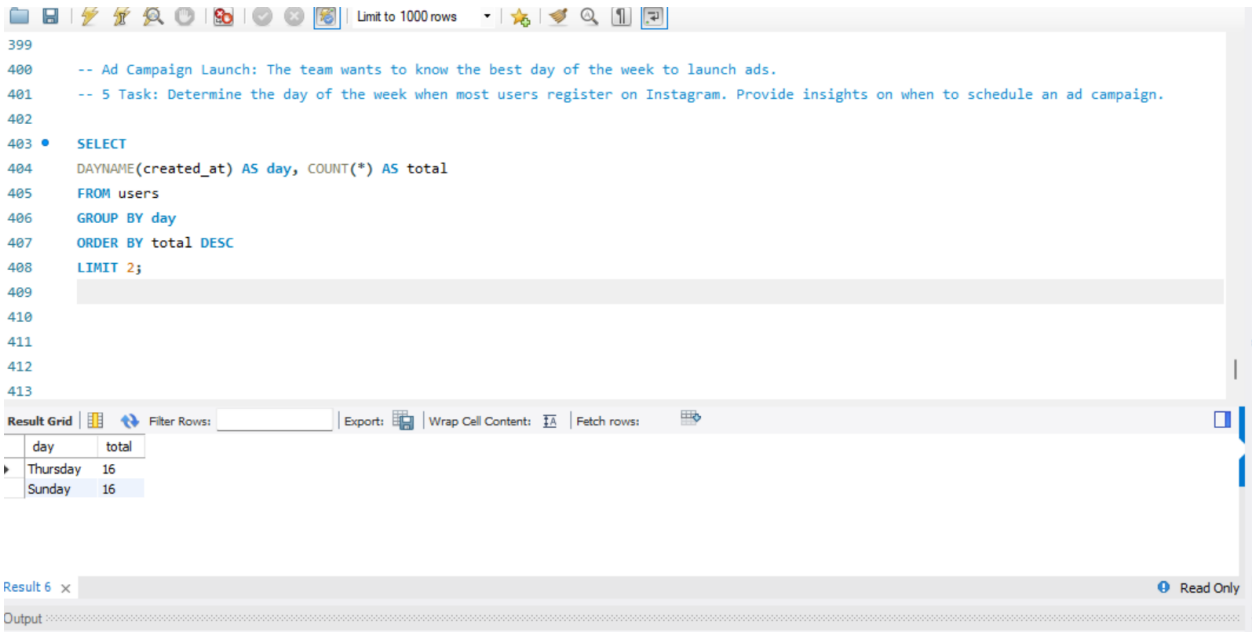


5. Ad Campaign Launch: The team wants to know the best day of the week to launch ads.
Your Task: Determine the day of the week when most users register on Instagram. Provide insights on when to schedule an ad campaign.

CODE:

```
SELECT

DAYNAME(created_at) AS day, COUNT(*) AS total

FROM users

GROUP BY day

ORDER BY total DESC

LIMIT 2;
```

OUTPUT:



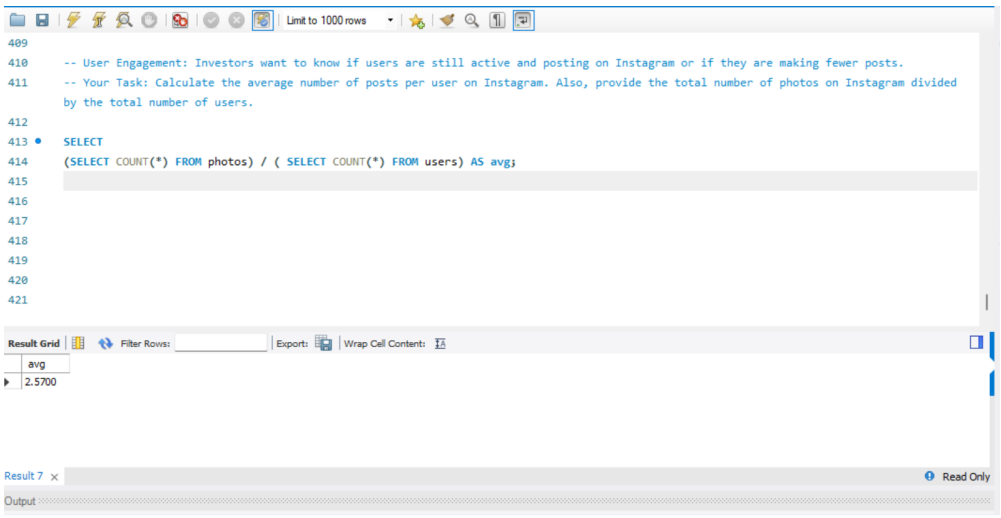
B) Investor Metrics:

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.

CODE:

```
SELECT

(SELECT COUNT(*) FROM photos) / ( SELECT COUNT(*) FROM users) AS avg;

--  Calculate the average number of posts per user on Instagram post count by user

select user_id, count(*)  as posts_count from photos

group by user_id

order by posts_count desc;

-- average_post per user

SELECT AVG(posts_count) as avg_posts_per_user FROM ( select user_id, count(*)  as posts_count from photos

group by user_id

order by posts_count desc) as user_posts;
```

OUTPUT:

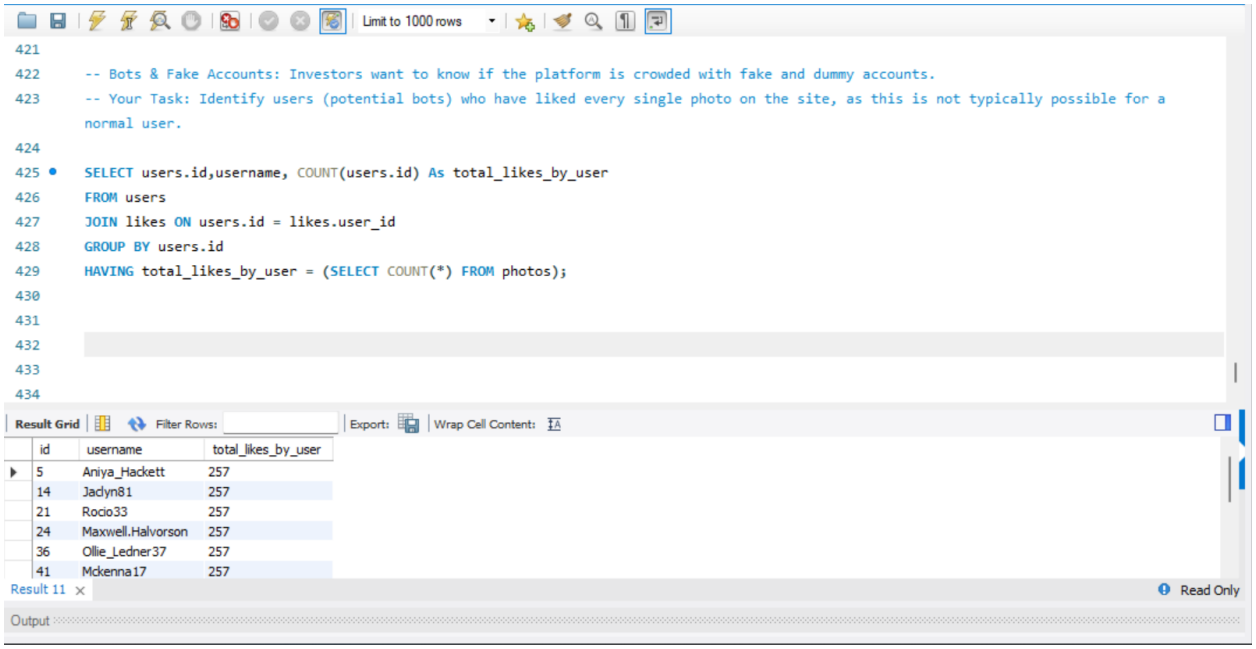


2. **Bots & Fake Accounts:** Investors want to know if the platform is crowded with fake and dummy accounts.
   Your Task: Identify users (potential bots) who have liked every single photo on the site, as this is not typically possible for a normal user.

CODE:

```
SELECT users.id,username, COUNT(users.id) As total_likes_by_user

FROM users

JOIN likes ON users.id = likes.user_id

GROUP BY users.id

HAVING total_likes_by_user = (SELECT COUNT(*) FROM photos);
```

OUTPUT:



Result

I learned from this assignment how data analysts and business professionals work with real-time data to make data-driven decisions. I gather that this project had a pretty modest dataset in terms of rows and columns, but even so, working on a project of this nature was a great experience. It greatly aided in my comprehension of the analysis process and offered guidance for making the best choice.