Instagram User Analytics

Project Description:

The project is about User Analysis which we can track how users engage and interact with Instagram User Analytics Dataset from Marketing Analysis in which marketing team wants to launch some campaigns as well as Investor Metrics where investors want to know if Instagram is performing well and is not becoming redundant like Facebook and perform analysis tasks per below.

1) Find out the oldest users who have been using Instagram from longest time.

2) As per Inactive Engagement program, find the users who haven't posted single photo yet on Instagram.

3) Under Contest Winner Declaration Program, Identify the user with most likes with single photo and provide data to team.

4) Identify and suggest the top 5 most commonly used hashtags on the platform.

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

6) Provide how many times does average user posts on Instagram. Also, provide the total number of photos on Instagram/total number of users.

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

Approach:

At first when I saw the data, it was Raw Data like unstructured data received in Doc file, where I unable to analyse data. So I used MYSQL Workbench, where I created database as well as all tables as per given in document to convert data in structured format. Then I started analysing the data with all tables one-by-one, where I got clear understanding of dataset. Saw the relationship between all tables by seeing Primary Keys & Foreign Keys among all columns, which are key player in relating the tables.

When it comes to perform the tasks, tried to understand the scenarios that are asked before performing tasks and then started executing the queries step by step to get desired outputs for all the tasks.

Tech-Stack Used:

I used MySQL Workbench-Version 8.0.34 software. It is hassle-free and easy to use RDBMS software. And queries were executed correctly. Also whenever got error it was clearly provided details of error with error code where I tried to diagnose the errors and corrected the queries and executed them.

Also MYSQL was advised by the team in the project dashboards, and guided by the mentor how to install and setup on system.

Insights:

Working on Instagram User Analytics have helped me to get insights on how to analyse the data and the relationship between the tables by seeing Primary Keys & Foreign Keys, As these key helped to Join the tables to fetch the correct records and based on that I conclude the functions or clauses like JOINS, SELECT, COUNT, GROUP BY, ORDER BY, HAVING, AS which are best suited while writing the queries.

Result:

Finally as a result I have achieved and gain knowledge on how to clean the data with the help of MySQL. And how to interact with database and how to customize the query to get the desired output. I learn about JOINS and fundamentals of SQL. How to analyse the given problem statement what are the functions we can use in SQL.

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.

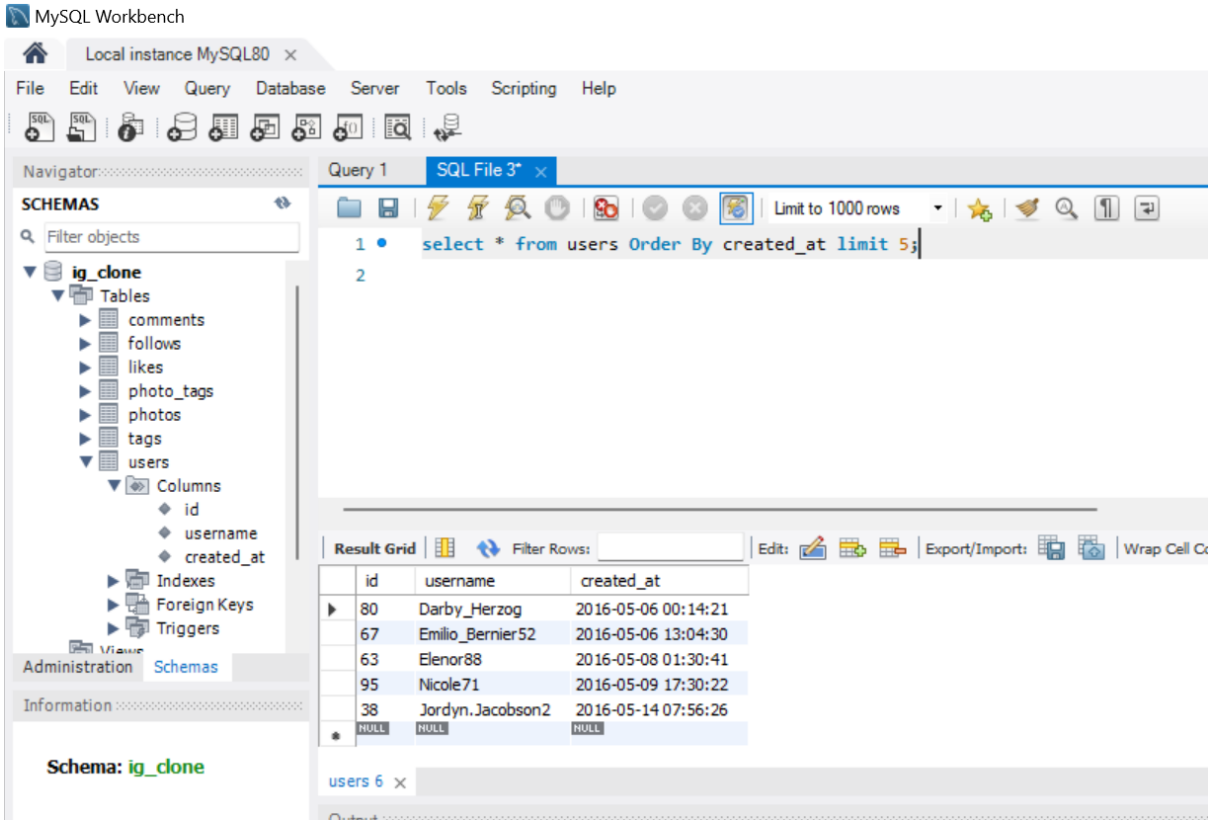
   Your Task: Identify the five oldest users on Instagram from the provided database.

Query:

```
select * from users Order By created_at limit 5;
```

Output:

	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



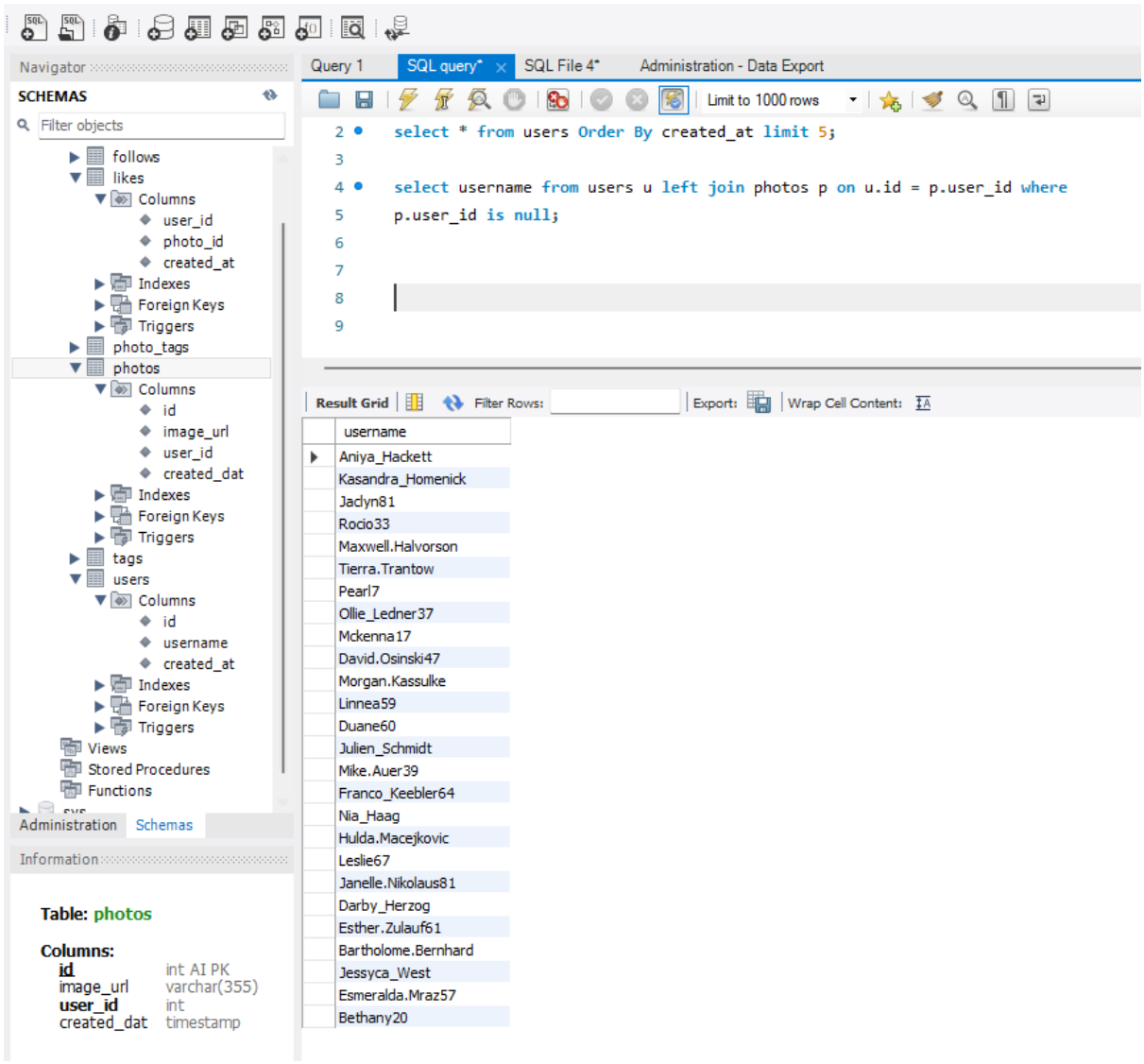
2. **Inactive User Engagement:** The team wants to encourage inactive users to start posting by sending them promotional emails.

   Your Task: Identify users who have never posted a single photo on Instagram.

Query:

```
select username from users u left join photos p on u.id = p.user_id where
p.user_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.

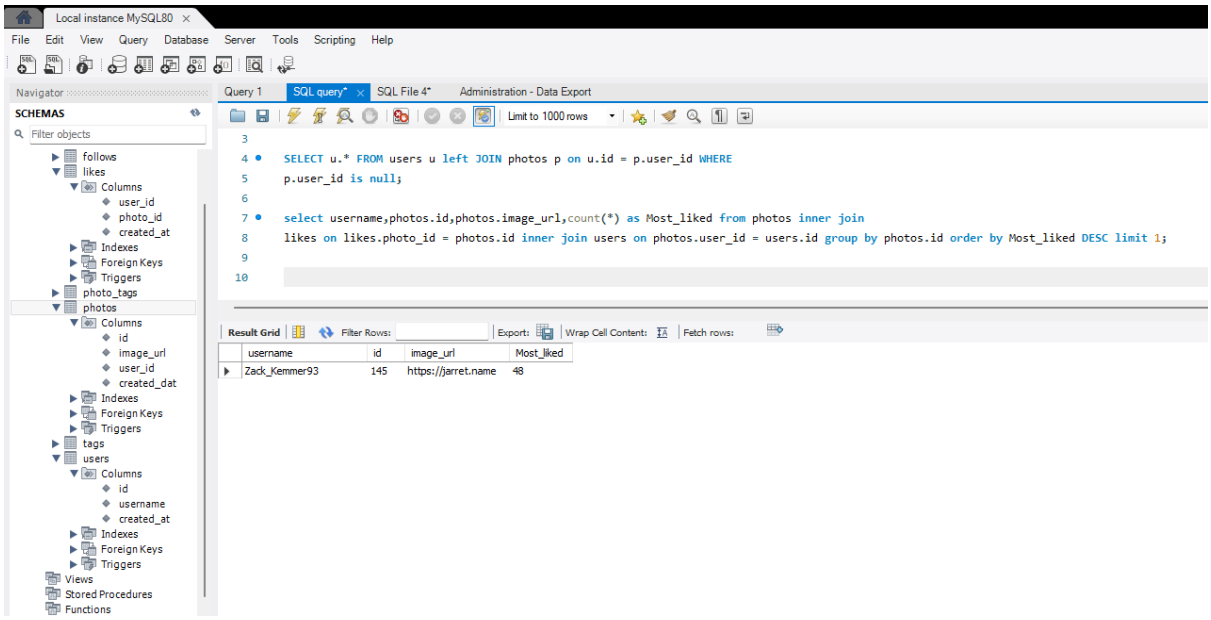
Your Task: Determine the winner of the contest and provide their details to the team.

Query:

```
select username,photos.id,photos.image_url,count(*) as Most_liked 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 Most_liked DESC limit 1;
```

Output:

	username	id	image_url	Most_liked
▶	Zack_Kemmer93	145	https://jarret.name	48



4. **Hashtag Research:** A partner brand wants to know the most popular hashtags to use in their posts to reach the most people.

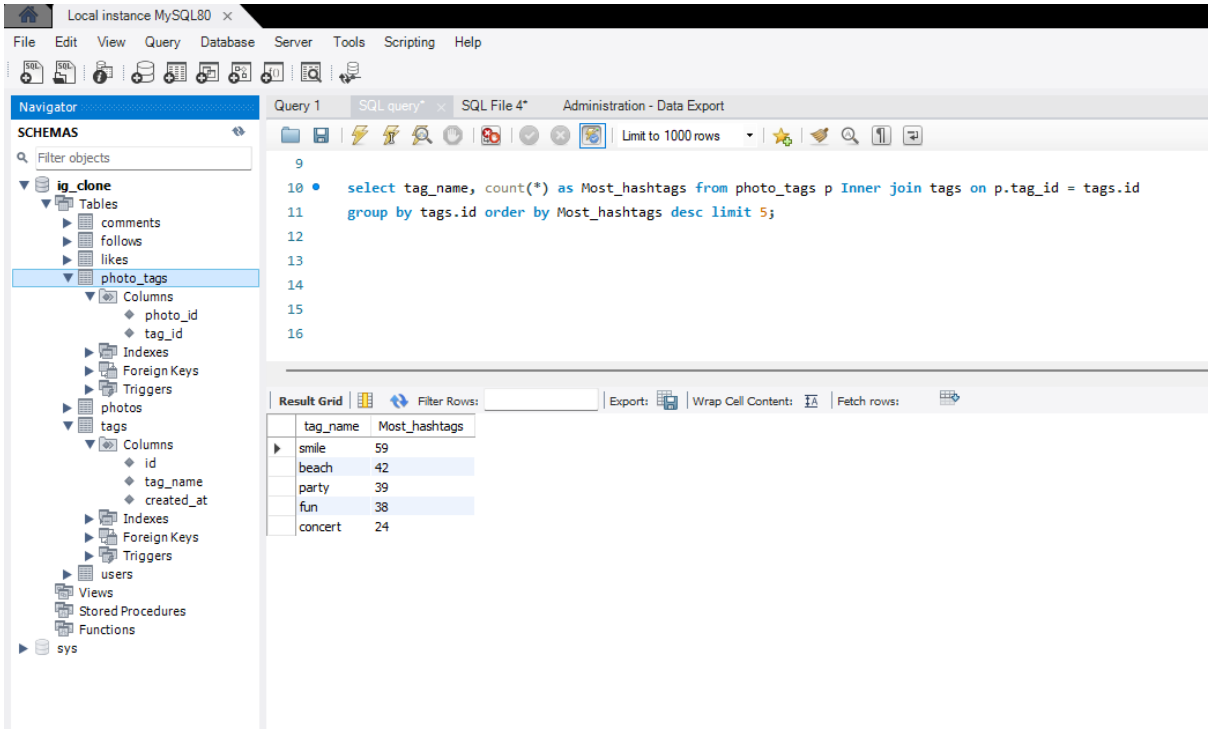
<u>Your Task: Identify and suggest the top five most commonly used hashtags on the platform</u>

Query:

```
select tag_name, count(*) as Most_hashtags from photo_tags p Inner join tags on p.tag_id = tags.id
group by tags.id order by Most_hashtags desc limit 5;
```

Output:

	tag_name	Most_hashtags
▶	smile	59
	beach	42
	party	39
	fun	38
	concert	24



5. **Ad Campaign Launch:** The team wants to know the best day of the week to launch ads.

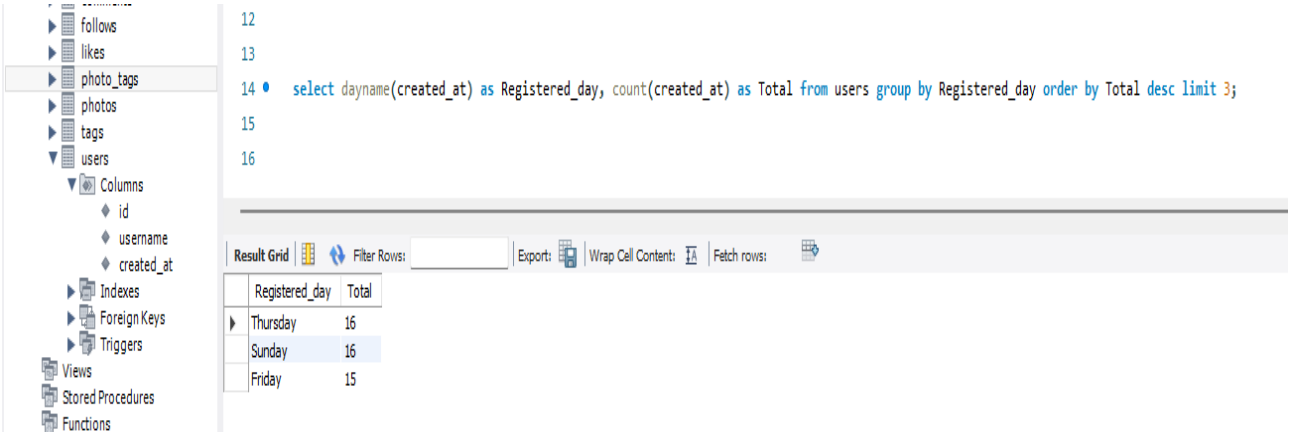
Your Task: Determine the day of the week when most users register on Instagram. Provide insights on when to schedule an ad campaign.

Query:

```
select dayname(created_at) as Registered_day, count(created_at) as Total from users
group by Registered_day order by Total desc limit 3;
```

Output:

	Registered_day	Total
▶	Thursday	16
	Sunday	16
	Friday	15



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.

i) Query:

```
/* Avg Post per user*/
Select count(*)/count(Distinct p.user_id) as AvgPostsPerUsers from photos p;
```

Output:

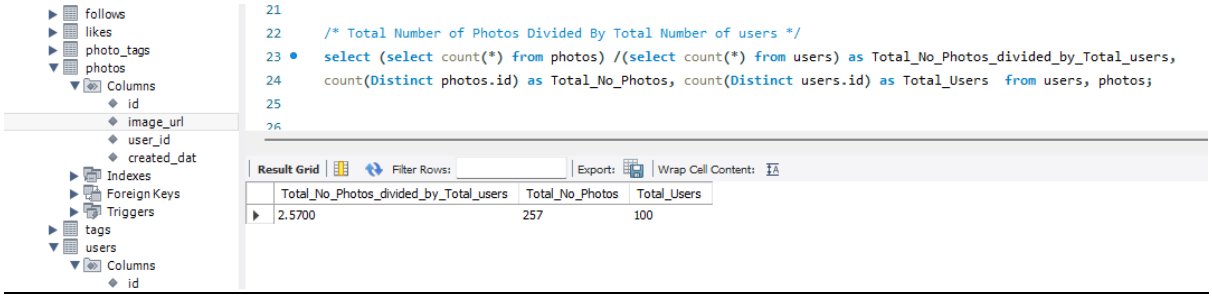
AvgPostsPerUsers
3.4730

ii) Query:

```
/* Total Number of Photos Divided By Total Number of users */
select (select count(*) from photos) /(select count(*) from users) as Total_No_Photos_divided_by_Total_users,
count(Distinct photos.id) as Total_No_Photos, count(Distinct users.id) as Total_Users  from users, photos;
```

Output:

Total_No_Photos_divided_by_Total_users	Total_No_Photos	Total_Users
2.5700	257	100



2. **Bots & Fake Accounts:** Investors want to know if the platform is crowded with fake and dummy accounts.

<u>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.</u>

Query:

```
select username as Potential_Bots, l.user_id as FakeIDs, count(l.photo_id) as Likes from likes l inner join users u on l.user_id = u.id
group by user_id having Likes = (select count(*) from photos);
```

Output:

Potential_Bots	FakeIDs	Likes
Aniya_Hackett	5	257
Jaclyn81	14	257
Rocio33	21	257
Maxwell.Halvorson	24	257
Ollie_Ledner37	36	257
Mckenna17	41	257
Duane60	54	257
Julien_Schmidt	57	257
Mike.Auer39	66	257
Nia_Haag	71	257
Leslie67	75	257
Janelle.Nikolaus81	76	257
Bethany20	91	257