

# Trainity Assignments

## Project 2 – Instagram User Analytics

### **Project Description:**

As per the project, I have been asked to analyze user interactions on Instagram user data and provide valuable insights to the product team regarding their questions, so I have to use SQL to work on the database provided and collect useful information for their marketing campaigns. A few questions have been asked by the management team regarding insights, which I have mentioned below:

### **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.

**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.

**3. Contest Winner Declaration:** The team has organized a contest where the user with the most likes on a single photo win.

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

**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.

**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.

### **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

**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.

### **Approach:**

For all the questions to be answered, I have installed a SQL workbench and later created a database. After getting all the information provided, I started writing queries for all the questions that had been raised and finally executed all the queries to get the output.

## Tech Stack used:

1. My SQL Workbench 8.0
2. Microsoft word Document

## Insights:

I learned a lot about SQL fundamentals from this project. While working on this project, I became aware of how simple it was to manage complex data and discovered the practical applications of SQL and databases.

## Result:

Given below are the SQL queries along with the output

### A) Marketing Analysis:

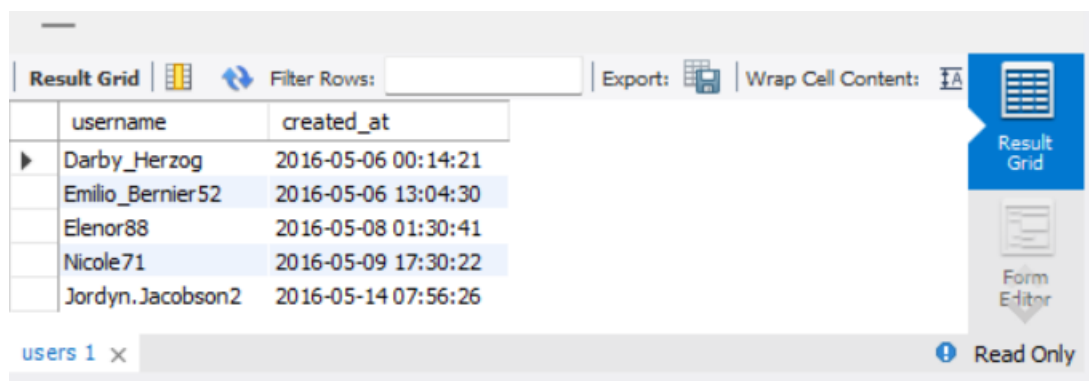
1. **Loyal User Reward:** The marketing team wants to reward the most loyal users, i.e., those who have been using the platform for the longest time.

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

#### **QUERY:**

```
use ig_clone;  
  
SELECT username, created_at  
  
FROM users  
  
ORDER BY created_at  
  
LIMIT 5;
```

#### **OUTPUT:**



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

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 u.username
FROM users u
LEFT JOIN photos p ON u.id=p.user_id
WHERE p.id IS NULL
ORDER BY u.username

```

#### OUTPUT:

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Result Grid | Form Editor | Read Only

username
Aniya_Hackett
Bartholome.Bernhard
Bethany20
Darby_Herzog
David.Osinski47
Duane60

Result 2 x

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Result Grid | Form Editor | Read Only

username
Esther.Zulauf61
Franco_Keebler64
Hulda.Macejkovic
Jadyn81
Janelle.Nikolaus81
Jessyca_West

Result 2 x

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Result Grid | Form Editor | Read Only

username
Hulda.Macejkovic
Jadyn81
Janelle.Nikolaus81
Jessyca_West
Julien_Schmidt
Kassandra_Homenick

Result 2 x

**3.Contest Winner Declaration:** The team has organized a contest where the user with the most likes on a single photo win.

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

#### QUERY:

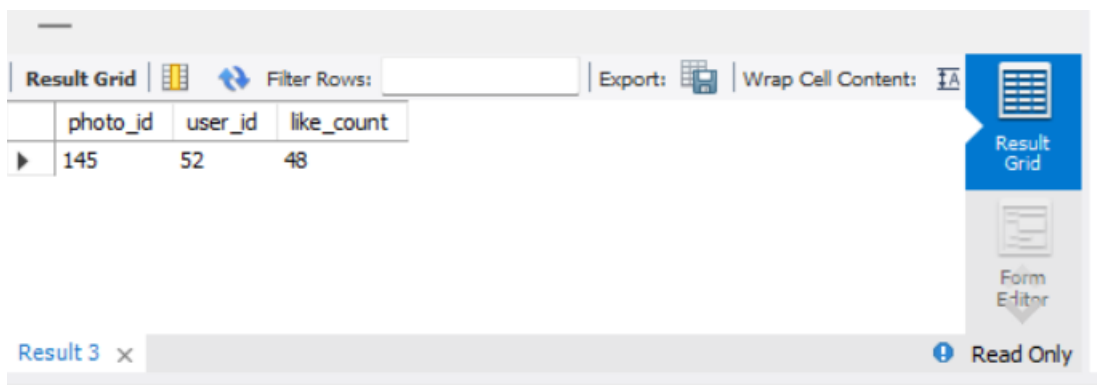
```

SELECT l.photo_id, p.user_id, COUNT(l.user_id) AS like_count
FROM likes l
JOIN photos p ON l.photo_id = p.id
GROUP BY l.photo_id

```

```
ORDER BY like_count DESC
LIMIT 1;
```

**OUTPUT:**



The screenshot shows a database interface with a 'Result Grid' tab. The grid contains one row of data with three columns: 'photo\_id', 'user\_id', and 'like\_count'. The values are 145, 52, and 48 respectively. The interface includes a 'Filter Rows' search bar, an 'Export' button, and a 'Wrap Cell Content' toggle. A 'Form Editor' button is also visible on the right side of the grid.

photo_id	user_id	like_count
145	52	48

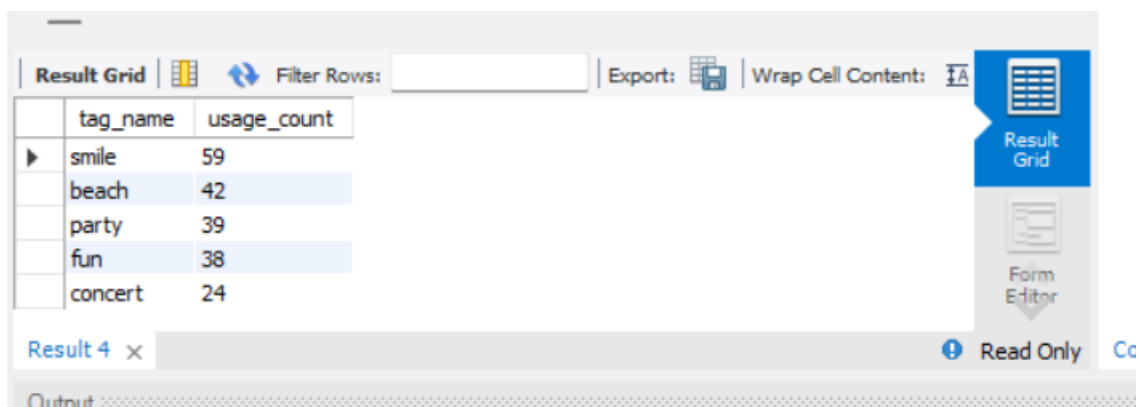
**4.Hashtag Research:** A partner brand wants to know the most popular hashtags to use in their posts to reach the most people.

Your Task: Identify and suggest the top five most commonly used hashtags on the platform.

**QUERY:**

```
SELECT t.tag_name,
COUNT(pt.photo_id) AS usage_count
FROM tags t
INNER JOIN photo_tags pt ON t.id = pt.tag_id
GROUP BY t.id, t.tag_name
ORDER BY usage_count DESC
LIMIT 5;
```

**OUTPUT:**



The screenshot shows a database interface with a 'Result Grid' tab. The grid contains five rows of data with two columns: 'tag\_name' and 'usage\_count'. The values are smile (59), beach (42), party (39), fun (38), and concert (24). The interface includes a 'Filter Rows' search bar, an 'Export' button, and a 'Wrap Cell Content' toggle. A 'Form Editor' button is also visible on the right side of the grid.

tag_name	usage_count
smile	59
beach	42
party	39
fun	38
concert	24

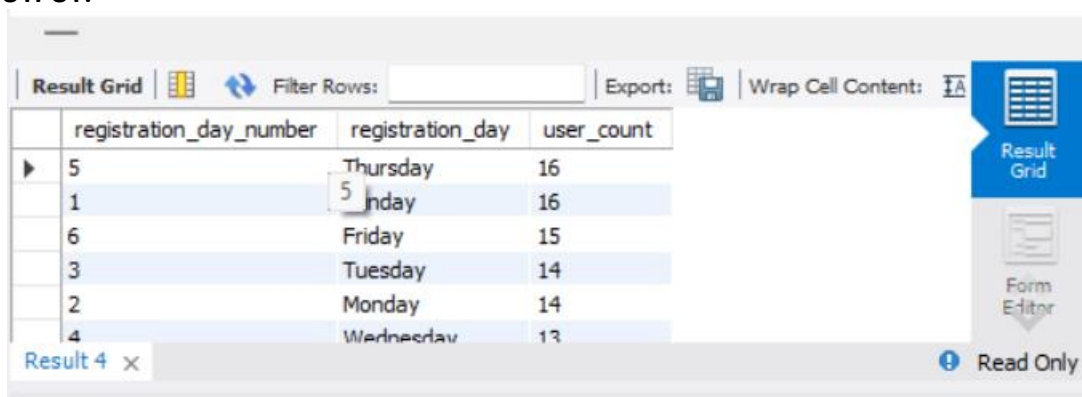
**5.Hashtag Research:** A partner brand wants to know the most popular hashtags to use in their posts to reach the most people.

Your Task: Identify and suggest the top five most commonly used hashtags on the platform.

**QUERY:**

```
SELECT DAYOFWEEK(created_at)AS registration_day_number,  
DAYNAME(created_at)AS registration_day,  
COUNT(id)AS user_count  
FROM users  
GROUP BY registration_day_number,registration_day  
ORDER BY user_count DESC;
```

**OUTPUT:**



The screenshot shows a 'Result Grid' interface with a table containing 5 rows of data. The columns are 'registration\_day\_number', 'registration\_day', and 'user\_count'. The data is sorted by 'user\_count' in descending order. The interface includes a 'Filter Rows' field, an 'Export' button, a 'Wrap Cell Content' toggle, and a 'Read Only' status indicator.

	registration_day_number	registration_day	user_count
▶	5	Thursday	16
	1	Monday	16
	6	Friday	15
	3	Tuesday	14
	2	Monday	14
	4	Wednesday	13

**B Investor Metrics:**

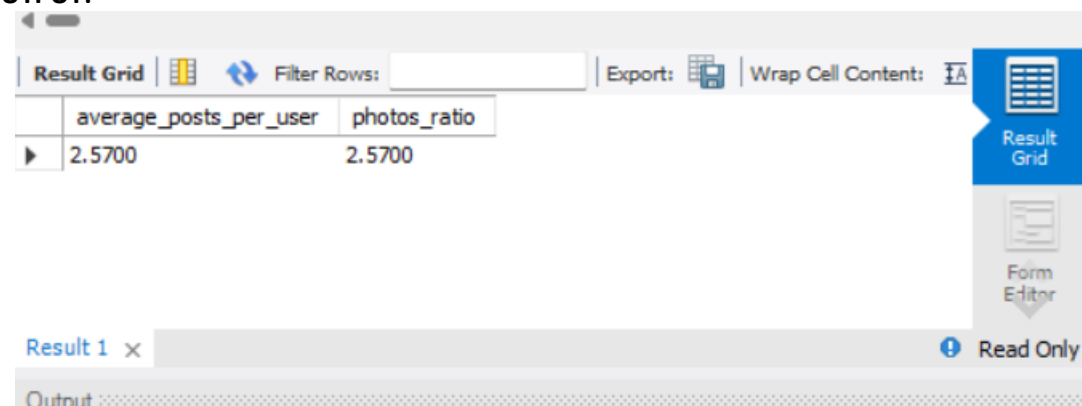
**1.User Engagement:** Investors want to know if users are still active and posting on Instagram or if they are making fewer posts.

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.

**QUERY:**

```
SELECT  
COUNT(p.id) / COUNT(DISTINCT u.id) AS average_posts_per_user,  
(SELECT COUNT(id) FROM photos) / (SELECT COUNT(id) FROM users) AS photos_ratio  
FROM users u  
LEFT JOIN photos p ON u.id = p.user_id;
```

**OUTPUT:**



The screenshot shows a 'Result Grid' interface with a table containing 2 rows of data. The columns are 'average\_posts\_per\_user' and 'photos\_ratio'. The data shows values of 2.5700 for both metrics. The interface includes a 'Filter Rows' field, an 'Export' button, a 'Wrap Cell Content' toggle, and a 'Read Only' status indicator.

	average_posts_per_user	photos_ratio
▶	2.5700	2.5700

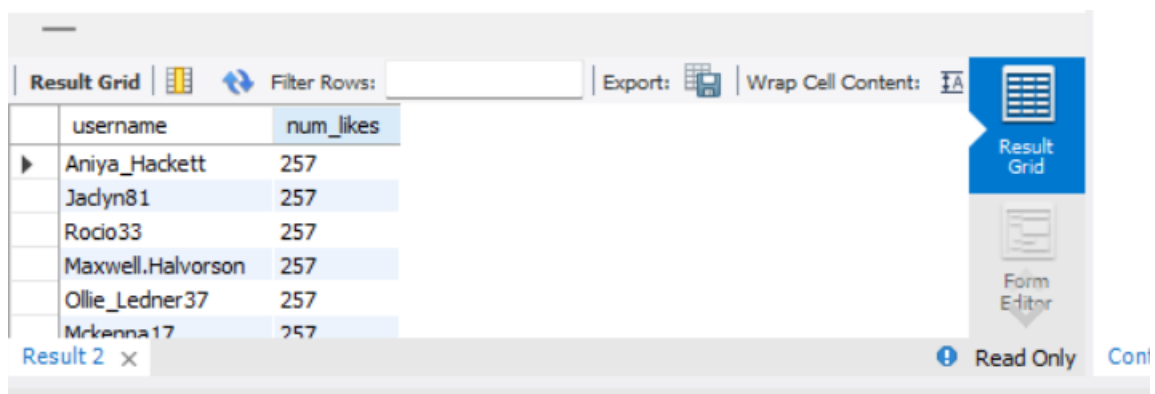
**2.Bots & Fake Accounts:** Investors want to know if the platform is crowded with fake and dummy accounts.

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

**QUERY:**

```
SELECT u.username, COUNT(l.photo_id) AS num_likes
FROM users u
JOIN likes l ON u.id = l.user_id
GROUP BY u.id
HAVING COUNT(l.photo_id) = (SELECT COUNT(*) FROM photos);
```

**OUTPUT:**



The screenshot shows a database interface with a 'Result Grid' tab selected. The grid displays the results of the SQL query, showing a list of usernames and their corresponding number of likes. The interface includes a toolbar with options like 'Filter Rows', 'Export', and 'Wrap Cell Content'. A 'Form Editor' button is also visible on the right side of the grid. The data is as follows:

username	num_likes
Aniya_Hackett	257
Jadyn81	257
Rocio33	257
Maxwell.Halvorson	257
Ollie_Ledner37	257
Mrkenna17	257

Finally, I analyzed and answered all the questions in the form of queries statements. Additionally, this project provided me with practical real- world experience and gave me good confidence about my SQL skills









