Instagram User Analytics Project Report

Analysis of User Engagement and Insights

T. SRI SATYA

Project Description:

This project aimed to analyze how users interact with the Instagram app and their level of engagement. By using SQL and MySQL Workbench, the goal was to uncover valuable insights that could help the marketing, product, and development teams make better decisions about the future of the app. This report summarizes the findings from the analysis of user data and provides recommendations based on the insights gathered.

Approach

1. **Database Creation:** A MySQL database was set up to store important user data, such as user profiles, photos, likes, and hashtags in the form of tables.
2. **Data Analysis:** SQL queries were run to gather insights about user engagement, marketing strategies, and metrics for potential investors. Each query focused on answering specific questions from the management team.
3. **Report Preparation:** The results were organized into this report, which includes snapshots of the SQL queries and their outputs to show the findings.

Tech-Stack Used

- **MySQL Workbench:** This tool was picked because it's easy to use and great for running SQL queries. It helps in managing the database and visualizing the results effectively.
- **MySQL:** This is the database system used to store and manage user data.

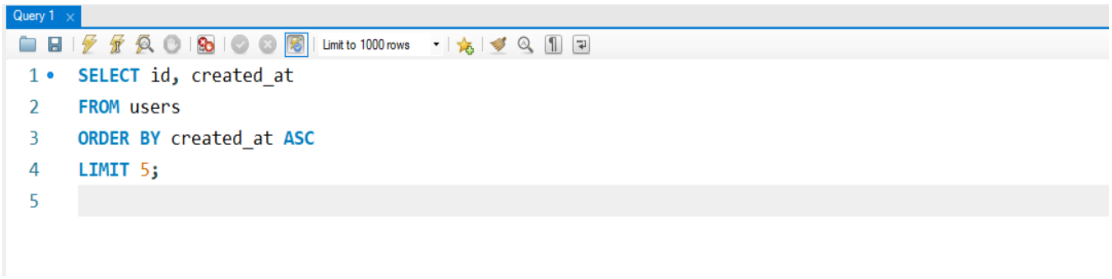
SQL Tasks and Insights

A) Marketing Analysis

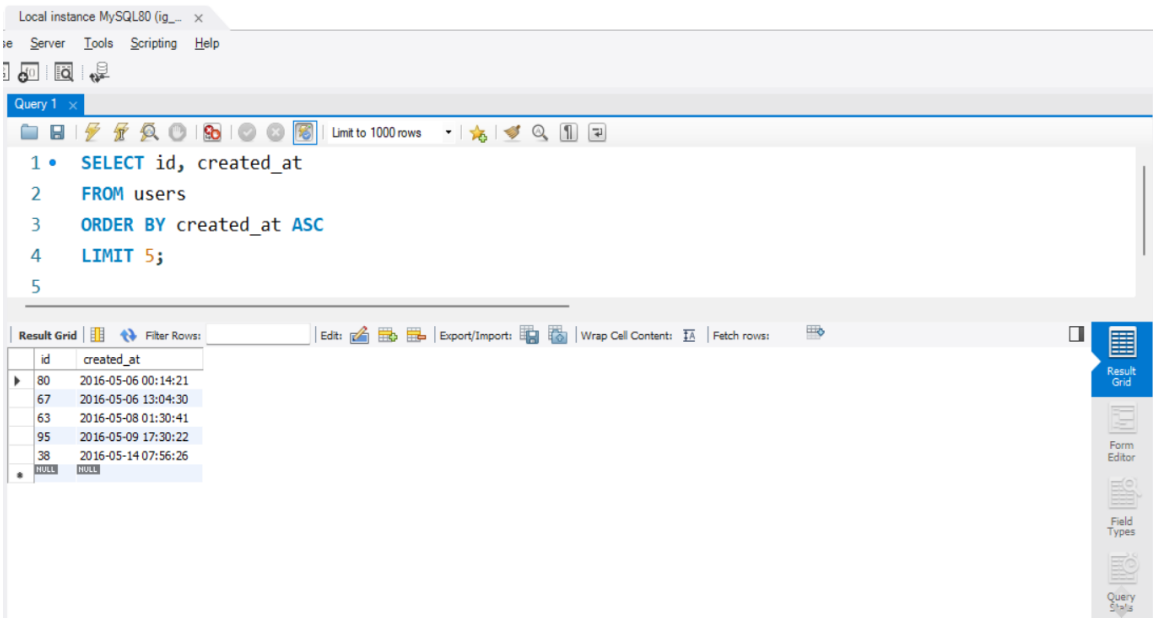
1. Loyal User Reward

Task: Identify the five oldest users on Instagram.

Query:



Output:



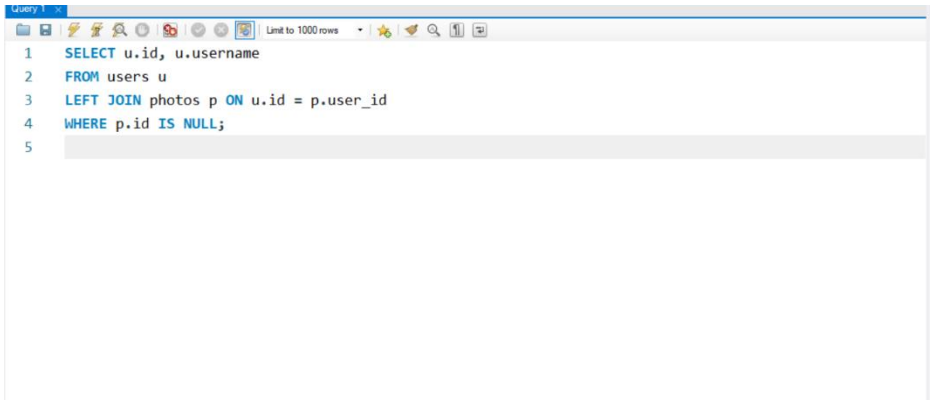
Insight:

The oldest users can be recognized for their long-term commitment, encouraging them to remain active on the platform. This could lead to higher engagement.

2. Inactive User Engagement

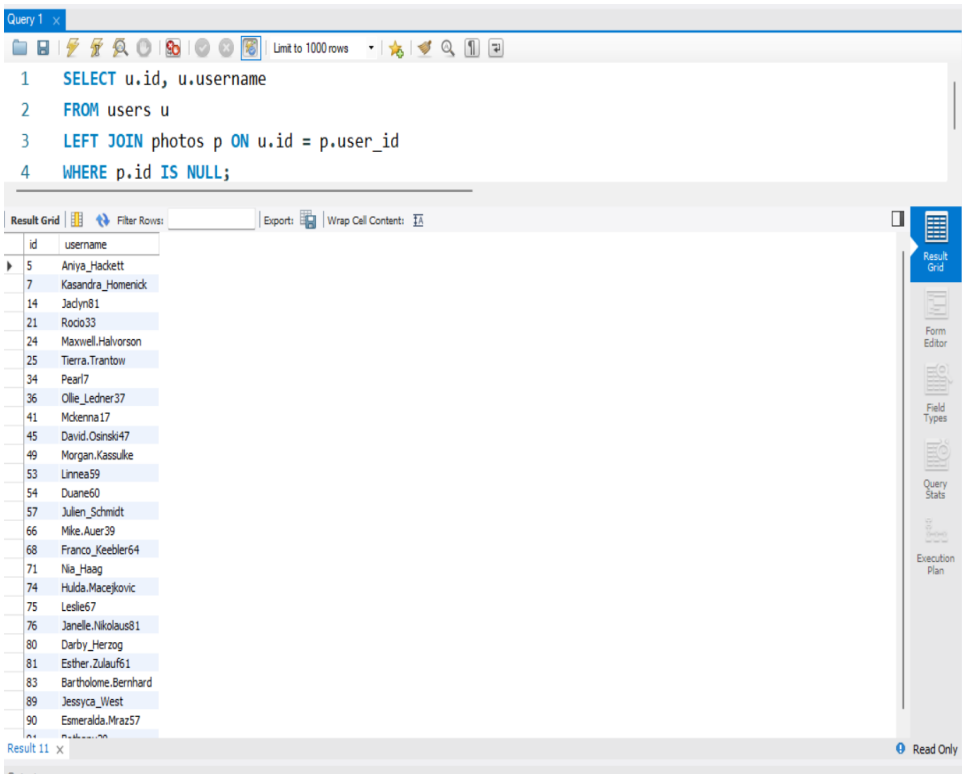
Task: Identify users who have never posted a single photo on Instagram.

Query:



```
SELECT u.id, u.username
FROM users u
LEFT JOIN photos p ON u.id = p.user_id
WHERE p.id IS NULL;
```

Output:



```
SELECT u.id, u.username
FROM users u
LEFT JOIN photos p ON u.id = p.user_id
WHERE p.id IS NULL;
```

id	username
5	Aniya_Hackett
7	Kasandra_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

Insight:

These users can be encouraged to post their first photo by sending them special offers or reminders, boosting their activity on the platform.

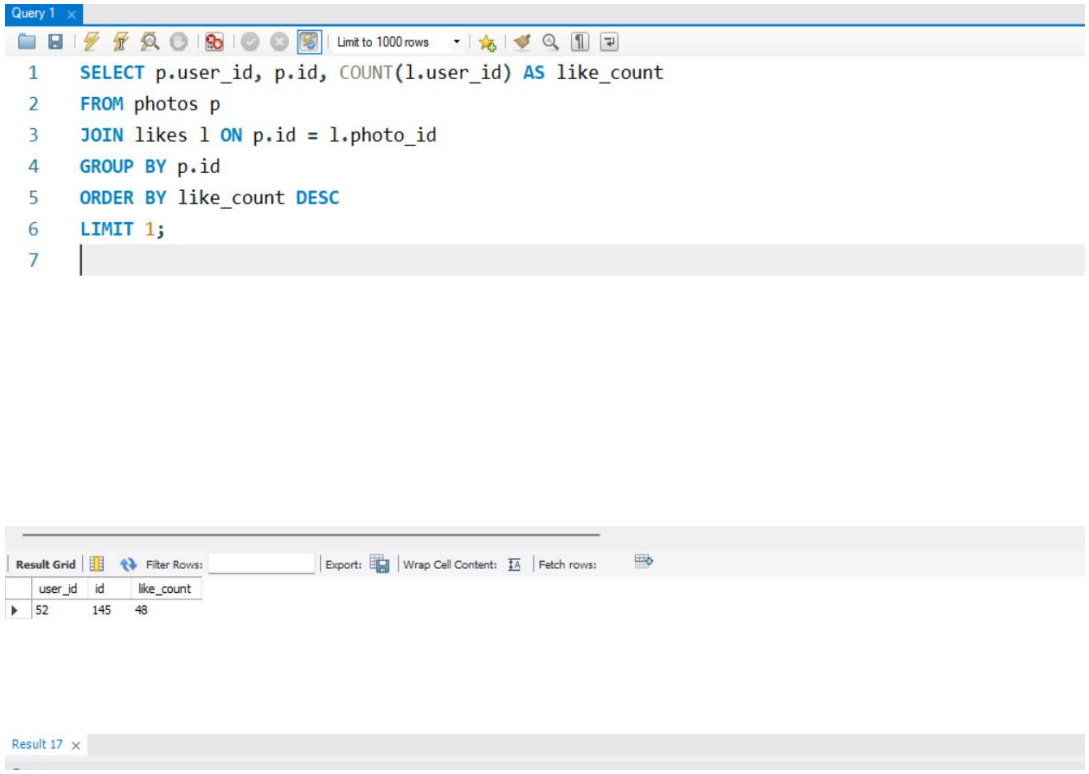
3. Contest Winner Declaration

Task: Determine the winner of the contest with the most likes on a single photo.

Query:

```
SELECT p.user_id, p.id, COUNT(l.user_id) AS like_count
FROM photos p
JOIN likes l ON p.id = l.photo_id
GROUP BY p.id
ORDER BY like_count DESC
LIMIT 1;
```

Output:



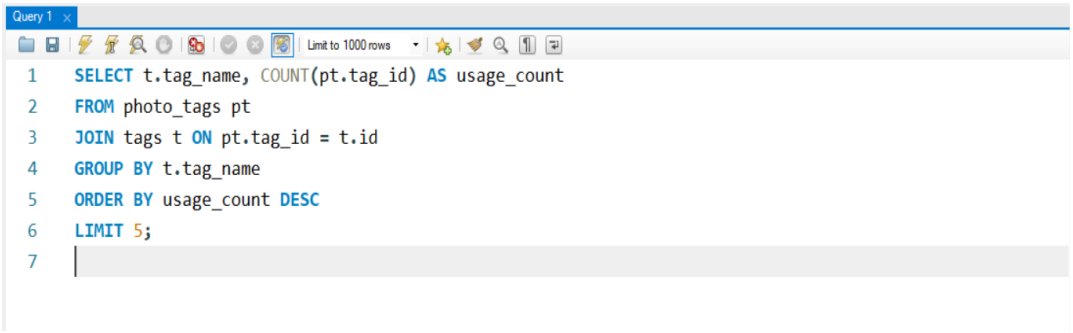
Insight:

The contest winner can be announced publicly, which will increase engagement and attract more users to future contests.

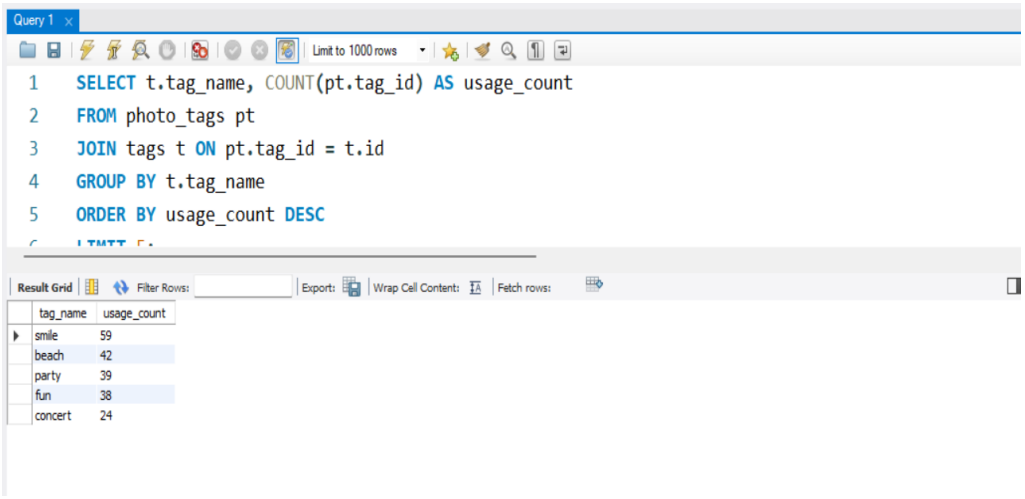
4. Hashtag Research

Task: Identify the top five most commonly used hashtags.

Query:



Output:



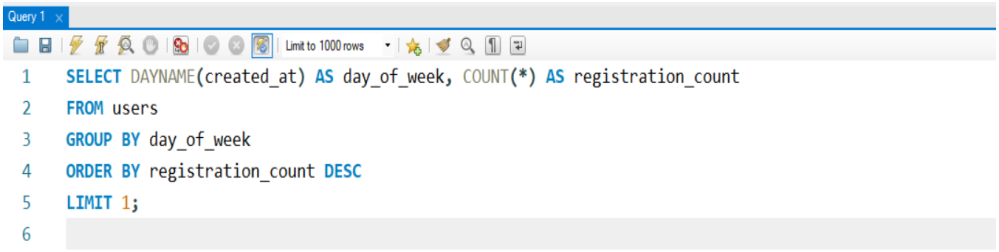
Insight:

These popular hashtags can help partners reach a wider audience and improve post visibility.

5. Ad Campaign Launch

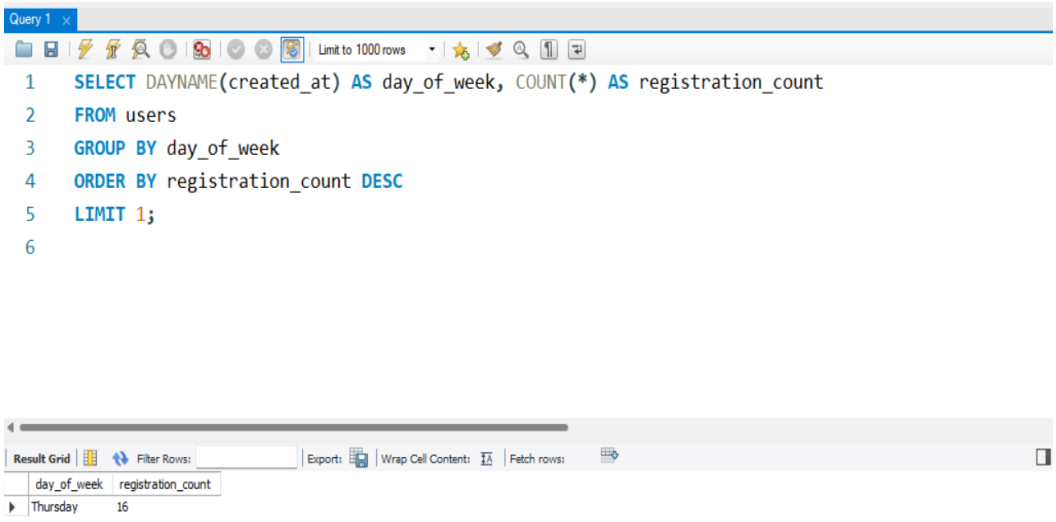
Task: Determine the best day of the week for launching ads.

Query:



```sql
SELECT DAYNAME(created_at) AS day_of_week, COUNT(*) AS registration_count
FROM users
GROUP BY day_of_week
ORDER BY registration_count DESC
LIMIT 1;
```

Output:



```sql
SELECT DAYNAME(created_at) AS day_of_week, COUNT(*) AS registration_count
FROM users
GROUP BY day_of_week
ORDER BY registration_count DESC
LIMIT 1;
```

day_of_week	registration_count
Thursday	16

Insight:

Running ads on the most popular day will likely increase engagement and reach more users.

B) Investor Metrics

1. User Engagement

Task: Calculate the average number of posts per user.

Query:

```
SELECT SUM(photo_count) AS total_no_of_photos,
       COUNT(user_id) AS total_no_of_users,
       SUM(photo_count) / COUNT(user_id) AS average_no_of_posts
FROM (
    SELECT users.id AS user_id, COUNT(photos.id) AS photo_count
    FROM users
    LEFT JOIN photos ON photos.user_id = users.id
    GROUP BY users.id
) AS photo_summary;
```

Output:

```
FROM (
    SELECT users.id AS user_id, COUNT(photos.id) AS photo_count
    FROM users
    LEFT JOIN photos ON photos.user_id = users.id
    GROUP BY users.id
) AS photo_summary;
```

total_no_of_photos	total_no_of_users	average_no_of_posts
257	100	2.5700

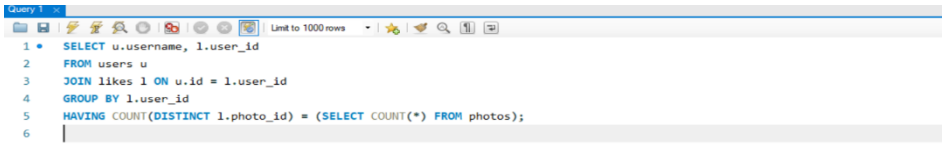
Insight:

The average number of posts per user indicates overall engagement on the platform. A higher average suggests active user participation, while a lower average may highlight the need for strategies to boost content creation.

2. Bots & Fake Accounts

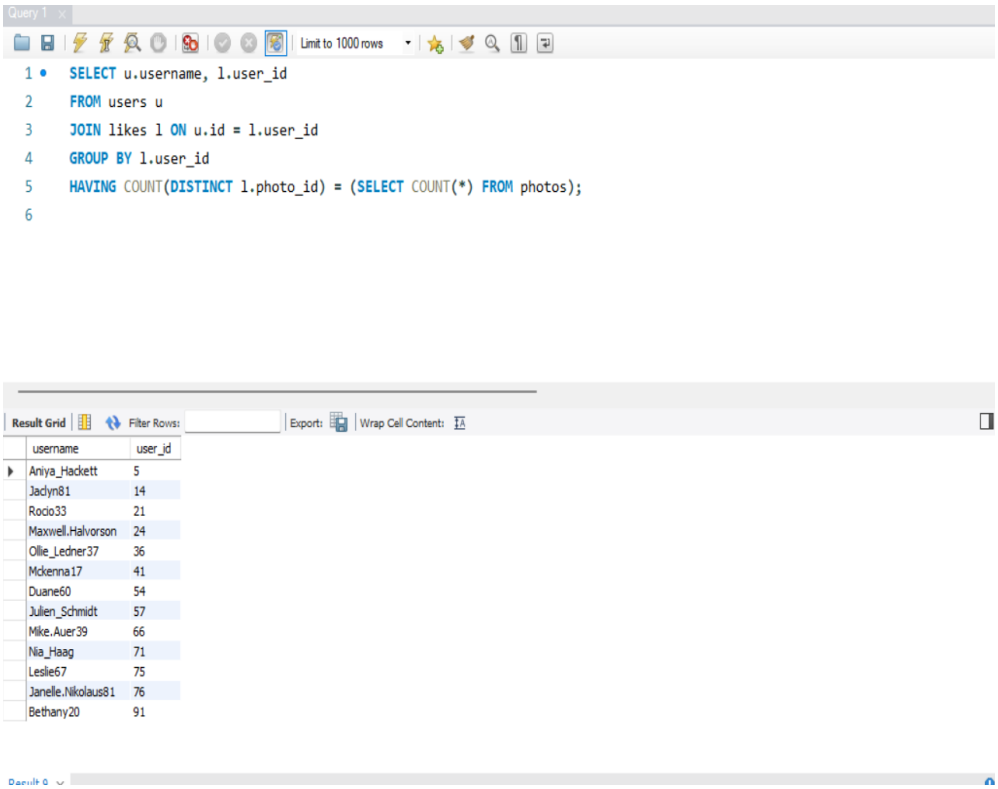
Task: Identify potential bots who have liked every single photo on the site.

Query:



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

Output:



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

username	user_id
Aniya_Hackett	5
Jadyn81	14
Rocio33	21
Maxwell.Halvorson	24
Ollie_Ledner37	36
Mckenna17	41
Duane60	54
Julien_Schmidt	57
Mike.Auer39	66
Nia_Haag	71
Leslie67	75
Janelle.Nikolaus81	76
Bethany20	91

Insight:

Identifying users who have liked every single photo can help flag potential bots or fake accounts. Removing them will improve the platform's trustworthiness.

Results:

1. Oldest Users: Identified the five oldest users on Instagram, which helps in recognizing loyal users for targeted rewards.
2. Inactive Users: Found users who have never posted a photo, indicating potential opportunities for re-engagement through promotional emails.
3. Contest Winner: Determined the user with the most likes on a single photo, who can be celebrated to enhance brand visibility and engagement.
4. Popular Hashtags: Identified the top five most commonly used hashtags, providing recommendations for partners to boost their posts' reach.
5. Best Day for Ads: Discovered the optimal day of the week for launching ads based on the highest user registrations, allowing for more effective ad scheduling.
6. Average User Engagement: Calculated the average number of posts per user, giving insight into user activity levels on the platform.
7. Potential Bots: Identified users who liked every photo, highlighting potential fake accounts that may need to be removed to maintain platform integrity.

Conclusion:

- The project successfully provided valuable insights into Instagram user engagement, helping the marketing and product teams make informed decisions.
- Recognizing loyal users and re-engaging inactive ones can enhance user retention and loyalty.
- Celebrating contest winners and promoting popular hashtags can increase brand visibility and user interaction.
- Understanding optimal ad launch days and average user engagement will help in better targeting and resource allocation for marketing campaigns.
- Identifying potential fake accounts ensures a healthier and more authentic user environment on Instagram.