

Name: Naman Rawal
Email: rawalnaman49@gmail.com

Instagram User Analytics

Analysis Overview:

This project involved analyzing Instagram user data to provide insights on various aspects such as user engagement, marketing strategies, and identifying fake accounts. The analysis was split into two main sections:

A. Marketing Insights:

1. Identifying Top Loyal Users:

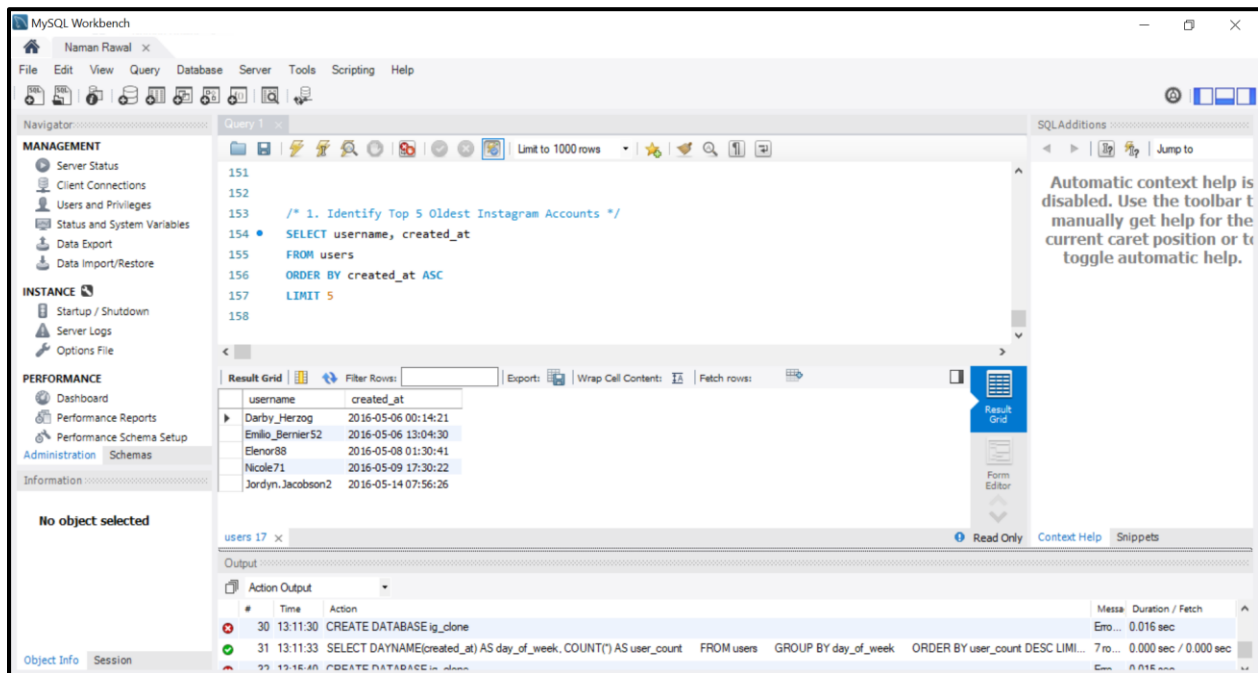
Focused on recognizing the platform's most consistent users (Top 5 oldest Instagram accounts).

- Approach:

Retrieved and sorted user data based on account creation date to identify the oldest users.

- SQL Query:

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



- Result:

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

2. Engaging Inactive Users:

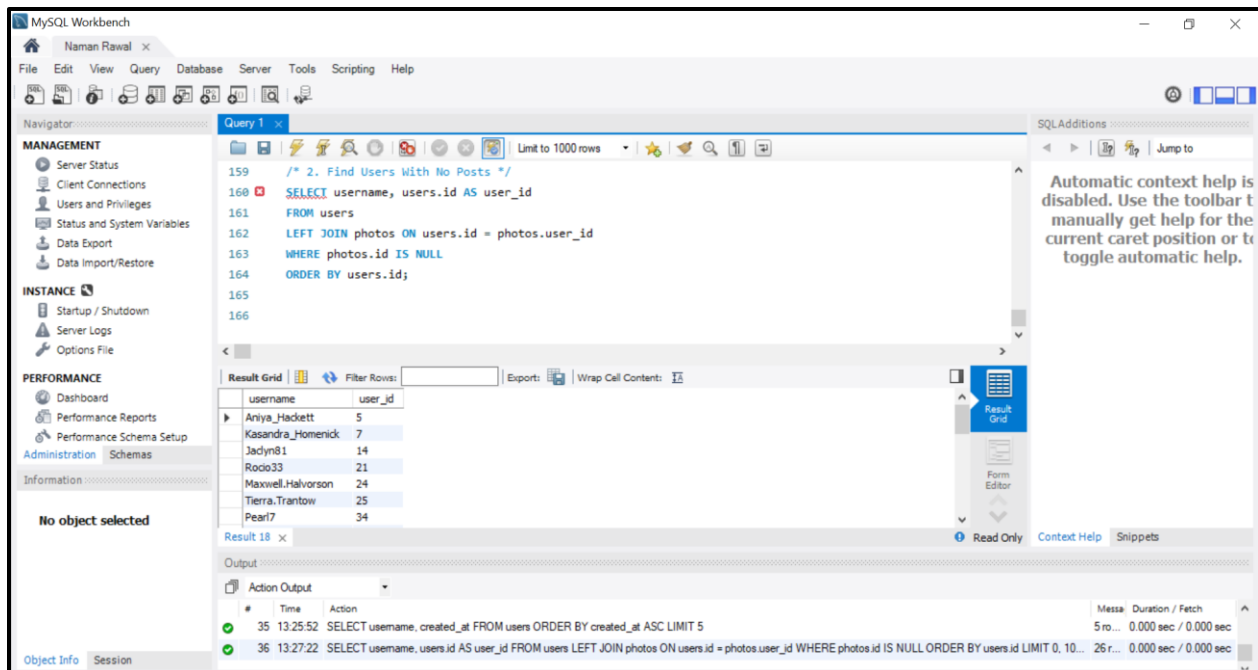
Targeted users who have never posted on the platform to encourage activity.

- Approach:

Joined user and photo data to identify users with no posts.

- SQL Query:

```
SELECT username, users.id AS user_id
FROM users
LEFT JOIN photos ON users.id = photos.user_id
WHERE photos.id IS NULL
ORDER BY users.id;
```



- **Result:**

username	user_id
Aniya_Hackett	5
Kasandra_Homenick	7
Jaclyn81	14
Rocio33	21
Maxwell.Halvorson	24
Tierra.Trantow	25
Pearl7	34
Ollie_Ledner37	36
Mckenna17	41
David.Osinski47	45
Morgan.Kassulke	49
Linnea59	53
Duane60	54
Julien_Schmidt	57
Mike.Auer39	66
Franco_Keebler64	68
Nia_Haag	71
Hulda.Macejkovic	74
Leslie67	75
Janelle.Nikolaus81	76
Darby_Herzog	80
Esther.Zulauf61	81
Bartholome.Bernhard	83
Jessyca_West	89
Esmeralda.Mraz57	90
Bethany20	91

3. Announcing Contest Winners:

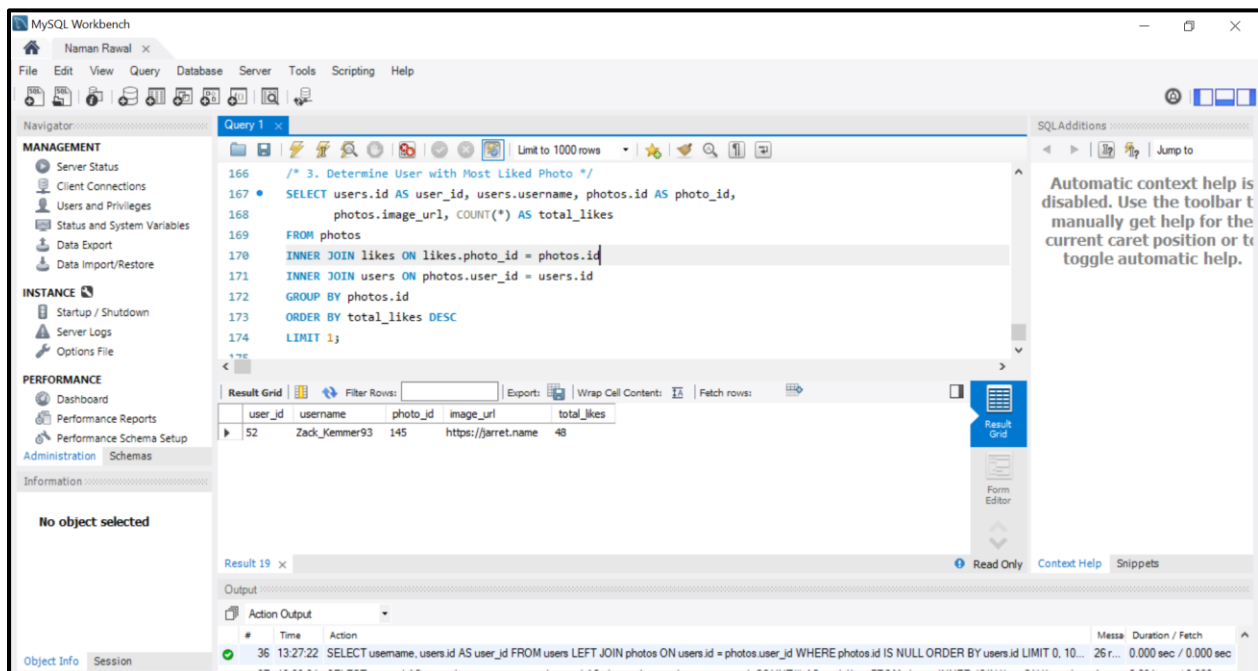
Determined the user with the most liked photo to announce contest winners.

- Approach:

Aggregated likes per photo to find the highest liked photo and its owner.

- SQL Query:

```
SELECT users.id AS user_id, users.username, photos.id AS photo_id,  
       photos.image_url, COUNT(*) AS total_likes  
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_likes DESC  
LIMIT 1;
```



The screenshot shows the MySQL Workbench interface. The 'Query' tab is active, displaying the following SQL query:

```
/* 3. Determine User with Most Liked Photo */  
SELECT users.id AS user_id, users.username, photos.id AS photo_id,  
       photos.image_url, COUNT(*) AS total_likes  
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_likes DESC  
LIMIT 1;
```

The 'Result Grid' tab shows the results of the query:

user_id	username	photo_id	image_url	total_likes
52	Zack_Kemmer93	145	https://jarret.name	48

The 'Output' tab shows the execution details:

#	Time	Action	Message	Duration / Fetch
36	13:27:22	SELECT username, users.id AS user_id FROM users LEFT JOIN photos ON users.id = photos.user_id WHERE photos.id IS NULL ORDER BY users.id LIMIT 0, 10...	26 r...	0.000 sec / 0.000 sec

- **Result:**

user_id	username	photo_id	image_url	total_likes
52	Zack_Kemmer93	145	https://jarret.name	48

4. Hashtag Analysis:

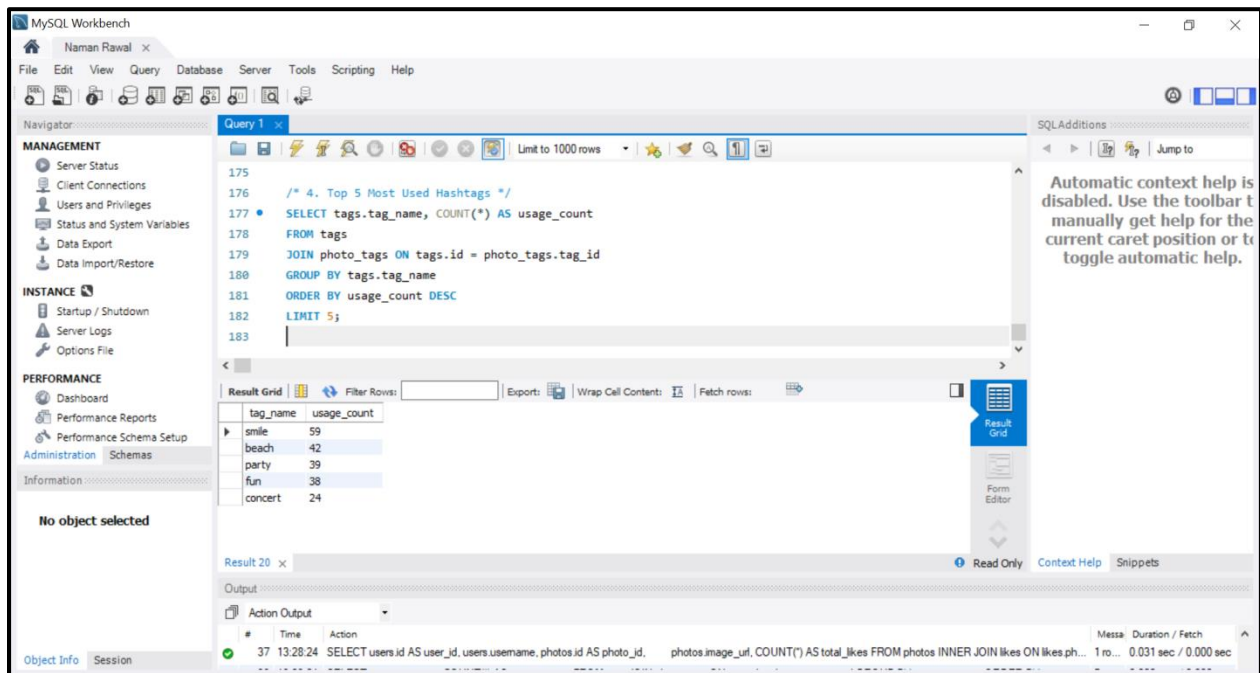
Evaluated the most commonly used hashtags to assist with targeted marketing.

- **Approach:**

Counted and ranked hashtags based on their usage frequency.

- **SQL Query:**

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



- Result:

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

5. Optimizing Ad Campaign Timing:

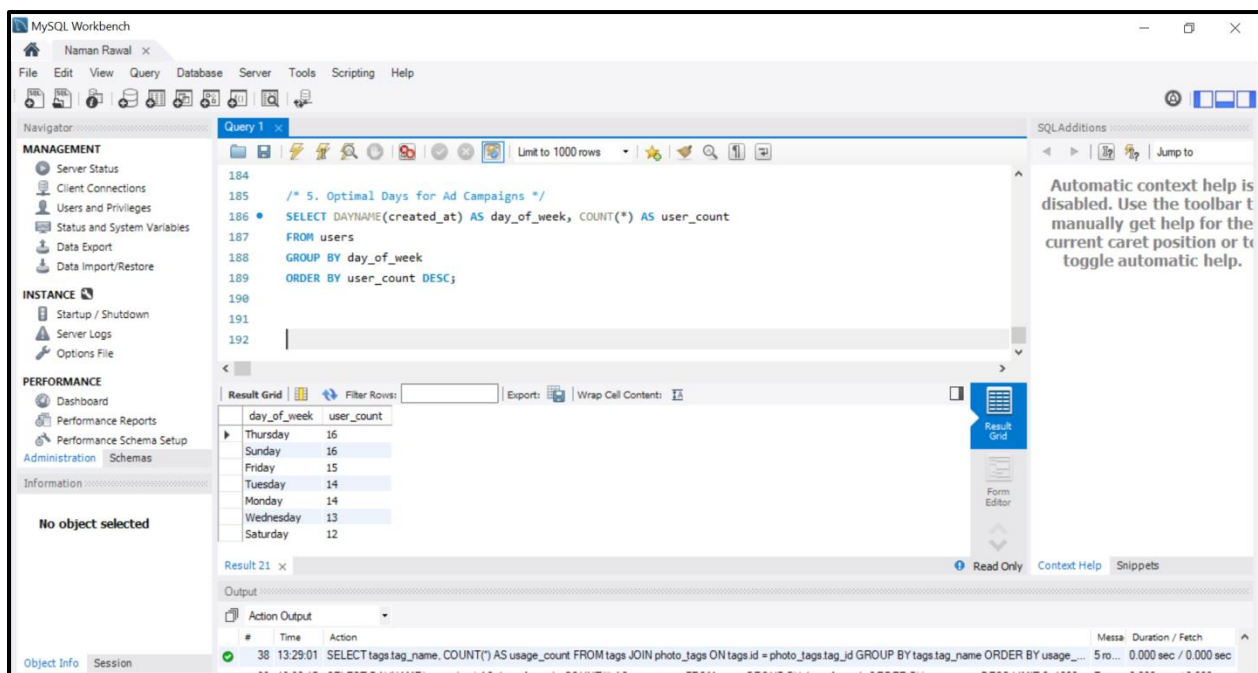
Suggested the best day to launch ads based on user registration trends.

- Approach:

Analyzed registration data to find the days with the highest user sign-ups.

- SQL Query:

```
SELECT DAYNAME(created_at) AS day_of_week, COUNT(*) AS user_count
FROM users
GROUP BY day_of_week
ORDER BY user_count DESC;
```



- **Result:**

day_of_week	user_count
Thursday	16
Sunday	16
Friday	15
Tuesday	14
Monday	14
Wednesday	13
Saturday	12

B. Investor Metrics:

1. User Engagement Metrics:

Assessed average user activity by analyzing post frequency.

- **Approach:**

Calculated average posts per user and total posts on the platform.

- **SQL Query:**

Here's how you can write the SQL queries to fulfill the task of calculating user engagement metrics:

1. Calculate the Average Number of Posts Per User:

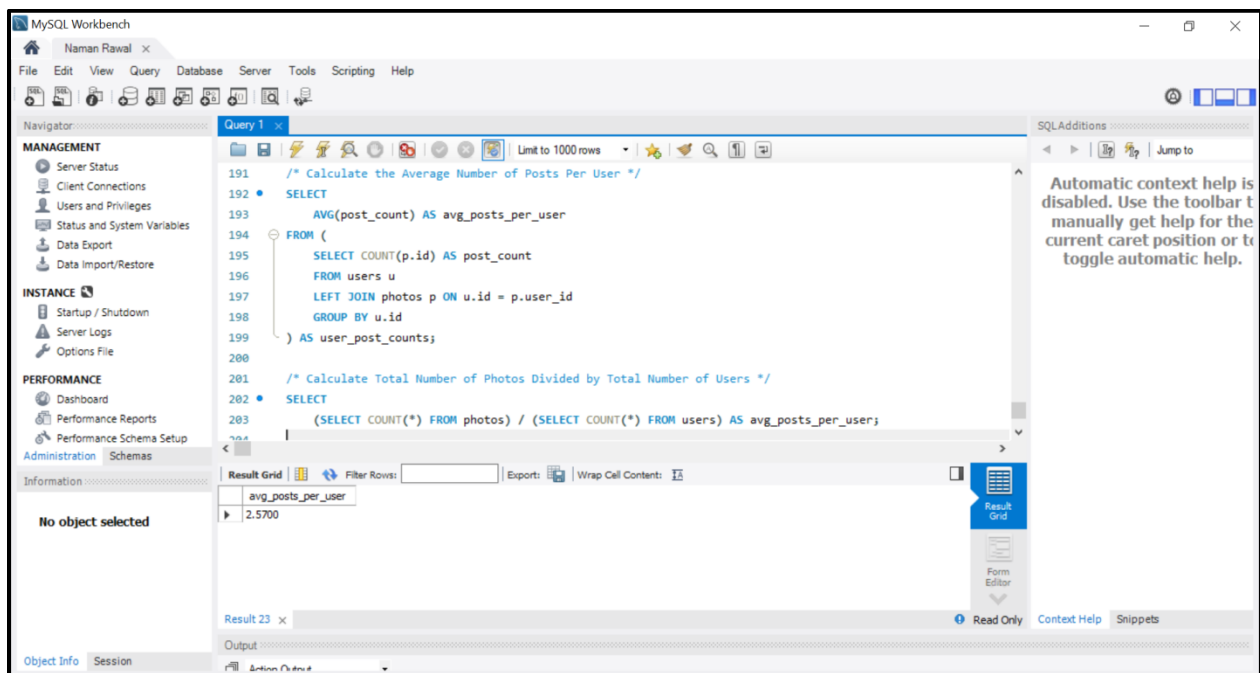
```
SELECT
  AVG(post_count) AS avg_posts_per_user
FROM (
  SELECT COUNT(p.id) AS post_count
  FROM users u
  LEFT JOIN photos p ON u.id = p.user_id
  GROUP BY u.id
) AS user_post_counts;
```

2. Provide the Total Number of Photos Divided by the Total Number of Users:

```
SELECT  
  (SELECT COUNT(*) FROM photos) / (SELECT COUNT(*) FROM users) AS  
avg_posts_per_user;
```

These queries will give:

- The average number of posts per user, considering all users.
- The ratio of the total number of photos to the total number of users, representing the average number of posts per user.



- **Result:**

avg_posts_per_user
2.57

2. Detecting Fake Accounts:

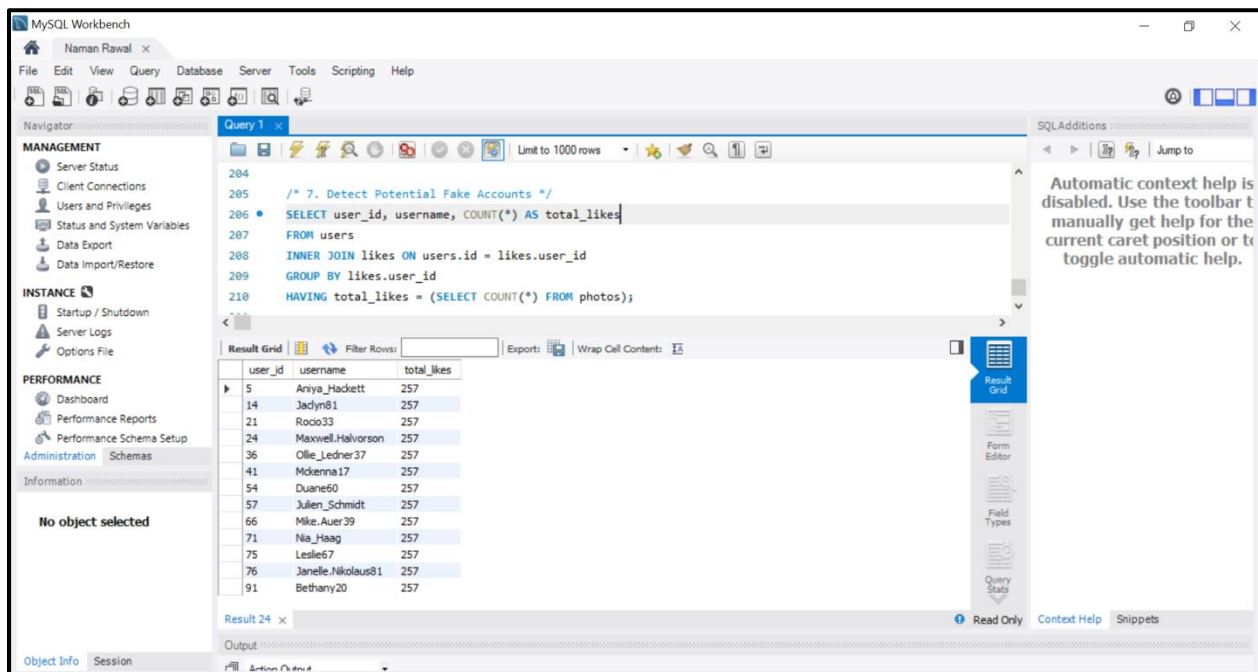
Identified potential fake accounts by analyzing users who liked every photo.

- Approach:

Cross-referenced user likes with total photos to flag suspicious behavior.

- SQL Query:

```
SELECT user_id, username, COUNT(*) AS total_likes
FROM users
INNER JOIN likes ON users.id = likes.user_id
GROUP BY likes.user_id
HAVING total_likes = (SELECT COUNT(*) FROM photos);
```



The screenshot shows the MySQL Workbench interface. The SQL editor contains the following query:

```
/* 7. Detect Potential Fake Accounts */
SELECT user_id, username, COUNT(*) AS total_likes
FROM users
INNER JOIN likes ON users.id = likes.user_id
GROUP BY likes.user_id
HAVING total_likes = (SELECT COUNT(*) FROM photos);
```

The query results are displayed in the Result Grid, showing a list of users and their total likes. The results are as follows:

user_id	username	total_likes
5	Aniya_Hackett	257
14	Jadyn81	257
21	Rocio33	257
24	Maxwell.Halvorson	257
36	Ollie_Ledner37	257
41	Mckenna17	257
54	Duane60	257
57	Julien_Schmidt	257
66	Mike_Auer39	257
71	Nia_Haag	257
75	Leslie57	257
76	Janelle.Nikolaus81	257
91	Bethany20	257

- Result:

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

Tools Utilized:

- **MySQL Workbench 8.0 CE:** Employed for executing SQL queries and performing the analysis.

Insights and Knowledge Gained

1. User Engagement Patterns:

- **Top Loyal Users:** Identifying the oldest accounts on Instagram highlighted the most long-standing and potentially loyal users. This information can be used to target these users for feedback or loyalty programs.

- **Inactive Users:** The list of users who have never posted provides a targeted group for re-engagement strategies. Encouraging these users to post could enhance overall engagement on the platform.

2. Marketing Strategies:

- **Contest Winners:** By identifying the photo with the highest number of likes, we can pinpoint influential users and leverage their popularity in marketing contests or campaigns.

- **Hashtag Analysis:** The most frequently used hashtags offer insights into current trends and user interests, which can help tailor marketing strategies to align with popular themes.

2. Optimization of Ad Campaigns:

- **Ad Campaign Timing:** Analyzing user registration trends revealed that Thursdays and Sundays are the most popular days for new sign-ups. Launching ad campaigns on these days may maximize visibility and effectiveness.

3. User Engagement Metrics:

- **Average Posts Per User:** The average number of posts per user (2.57) indicates a moderate level of user activity. This metric helps in understanding overall engagement and can guide strategies to increase user interaction.

- **Photo-to-User Ratio:** The ratio of total photos to users supports the average posts per user calculation and provides a benchmark for user activity levels.

4. Identifying Potential Fake Accounts:

- **Fake Accounts Detection:** Users who liked every photo are flagged as potential fake accounts. This insight helps in maintaining the integrity of user data and identifying suspicious activity.

Summary:

The project provided a comprehensive analysis of Instagram user data, uncovering trends in user engagement, marketing effectiveness, and account authenticity. Key findings include the identification of loyal users, strategies for engaging inactive accounts, and optimization of marketing campaigns based on user activity patterns. These insights are valuable for enhancing user engagement and refining marketing strategies on Instagram.