MARKETING ANALYSIS

The five oldest users on Instagram

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

Users who have never posted a single photo on Instagram

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

Find the user with the most likes on a single photo

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

The top five most commonly used hashtags on the platform

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

The day of the week when most users register on Instagram

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

INVESTOR METRICS

The average number of posts per user on Instagram. Also,the total number of photos on Instagram divided by the total number of users

```
SELECT AVG(photo_count) AS avg_posts_per_user
FROM (
    SELECT user_id, COUNT(*) AS photo_count
    FROM photos
    GROUP BY user_id
) AS user_photo_counts;

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

No.of users(potential bots) who have liked every single photo on the site

```
SELECT l.user_id
FROM likes l
GROUP BY l.user_id
HAVING COUNT(l.photo_id) = (SELECT COUNT(*) FROM photos);
```

OR

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

);
```

Project Description

This project involves analyzing user interaction and engagement data from Instagram to derive insights about user behavior, content popularity, and potential anomalies .

Approach

1. **Data Understanding**:
   - Reviewed the database schema and tables: `users`, `photos`, `comments`, `likes`, `follows`, `tags`, and `photo_tags`.
   - Identified the relationships between tables and the key columns necessary for analysis.
2. **Data Analysis**:
   - **User Activity Analysis**: Calculated the average number of posts per user and identified users who have never posted a photo.
   - **Content Popularity**: Determined the top five most commonly used hashtags.
   - **User Registration Trends**: Identified the day of the week with the highest number of user registrations.
   - **Anomaly Detection**: Identified users who have liked every single photo, indicating potential bot activity.

Tech-Stack Used

1. **DB Fiddle: MY SQL V8.0**

Insights

1. **Average Posts per User**:
   - Calculated the average number of posts per user to understand user engagement levels.
   - Found that on average, each user posts a certain number of photos, indicating active participation on the platform.
2. **Top Hashtags**:
   - Identified the top five most commonly used hashtags, revealing trending topics and popular content themes.
3. **User Registration Trends**:
   - Determined the day of the week with the highest number of registrations, suggesting optimal times for user engagement and potential ad campaign scheduling.
4. **Anomaly Detection**:
   - Identified users who have liked every single photo, flagging them as potential bots for further investigation.

Result

1. **Achievements**:
   - Successfully analyzed user interaction and engagement data to provide actionable insights.
   - Identified key patterns in user behavior, content popularity, and potential anomalies.
2. **Impact**:
   - The analysis helped in understanding user activity and engagement patterns, aiding in optimizing platform features and user experience.
   - Provided valuable information for scheduling ad campaigns to maximize reach and effectiveness.
   - Identified potential bot activity, contributing to maintaining platform integrity and security.