project description-
This project focuses on analyzing Instagram user data to provide actionable insights that can help drive strategic decisions for user engagement, marketing efforts, and platform growth. Using SQL and MySQL Workbench, I performed tasks such as identifying the most loyal users, detecting inactive accounts, determining popular hashtags, and calculating user engagement metrics. These analyses aim to assist the product, marketing, and development teams in optimizing the platform's features and improving overall user experience, while also addressing concerns about potential bot activity and identifying opportunities for targeted ad campaigns.

Approach-
The approach to this project involved a structured analysis of Instagram's user data using SQL queries. The following steps were taken:

1. Database Setup: I started by setting up the provided database in MySQL Workbench. This database contained tables with information about users, posts, likes, and hashtags.
2. Task-Specific Queries: For each question posed by the management team, I designed and executed SQL queries tailored to extract the necessary data. The queries were focused on identifying the oldest users, inactive users, contest winners, popular hashtags, and registration patterns.
3. Data Analysis: I analyzed the data retrieved through SQL to understand patterns and trends, such as the behavior of active versus inactive users, and days of the week with the highest registrations.
4. Result Documentation: Screenshots of the queries and outputs were captured and integrated into the final report. The insights derived were documented in a concise and relevant manner to support business decision-making.

Tech-Stack used-

MySQL Workbench (Version 8.0): Chosen for its robust capabilities in managing, querying, and visualizing relational databases. It allows for efficient database querying and supports complex SQL queries.

SQL (Structured Query Language): Used to interact with the database, perform data extraction, and analyze the datasets to answer business questions.

Insights-

Throughout the analysis, several key insights emerged:

- Loyal Users: The oldest users on Instagram were identified, which will assist the marketing team in rewarding long-term loyalty and increasing engagement.
- Inactive Users: A considerable number of users had never posted any photos. This indicates a potential area for marketing outreach to encourage these users to become more active on the platform.

- Popular Hashtags: The top five most used hashtags were identified, revealing patterns in user preferences and brand reach. This insight can assist partners in targeting posts for greater visibility.
- Optimal Ad Campaign Day: The analysis showed that the majority of users registered on certain days of the week, indicating the best times to schedule ad campaigns for maximum exposure.
- User Engagement: The average number of posts per user was calculated, providing insights into how actively engaged Instagram's user base is.
- Bot Detection: Potential bots were identified by finding users who liked every post on the platform—behavior that is typically unrealistic for genuine users. This helps the platform maintain authenticity by addressing fake accounts.

These insights, derived from the SQL queries, provide a foundation for making strategic decisions that can positively impact user engagement, marketing efforts, and overall platform growth.

Result-

Through the SQL-based analysis of Instagram's user data, several key results were achieved, providing valuable insights for business decision-making:

1. Loyal User Identification: I successfully identified the five oldest users on Instagram based on their registration dates, allowing the marketing team to focus on rewarding user loyalty and engagement.

```
94
95 •  SELECT id, username, created_at
96     FROM users
97     ORDER BY created_at ASC
98     LIMIT 5;
99
```

Result Grid | Filter Rows: | Edit: | Export/I

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 analysis revealed a list of users who had never posted any photos, helping the marketing team to target them for re-engagement efforts through promotional emails.

```
99
100 •  SELECT users.id, users.username
101    FROM users
102    LEFT JOIN photos  ON users.id = photos.user_id
103    WHERE photos.user_id IS NULL;
104
```

id	username
5	Aniya_Hackett
7	Kasandra_Homenick
14	Jaclyn81
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

3. Contest Winner: I determined the contest winner by identifying the user with the most likes on a single photo, providing the required user details for the team to announce the winner.

```
104
105 •  SELECT u.id AS user_id, u.username, p.id AS photo_id, COUNT(l.photo_id) AS like_count
106    FROM photos p
107    JOIN users u ON p.user_id = u.id
108    JOIN likes l ON p.id = l.photo_id
109    GROUP BY p.id, u.id, u.username
110    ORDER BY like_count DESC
111    LIMIT 1;
112
113
114
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

user_id	username	photo_id	like_count
52	Zack_Kemmer93	145	48

4. Hashtag Popularity: The top five most frequently used hashtags were identified, offering valuable information to partner brands about which hashtags to use for broader reach and engagement.

```
113 •   SELECT t.tag_name, COUNT(pt.tag_id) AS usage_count
114     FROM tags t
115     JOIN photo_tags pt ON t.id = pt.tag_id
116     GROUP BY t.tag_name
117     ORDER BY usage_count DESC
118     LIMIT 5;
119
120
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

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

5. Optimal Ad Campaign Timing: The analysis pinpointed the best day of the week for launching ad campaigns based on user registration data, offering insights into when ad campaigns are likely to reach the most users.

```
120 •   SELECT DAYNAME(created_at) AS registration_day, COUNT(id) AS user_count
121     FROM users
122     GROUP BY registration_day
123     ORDER BY user_count DESC
124     LIMIT 1;
125
```

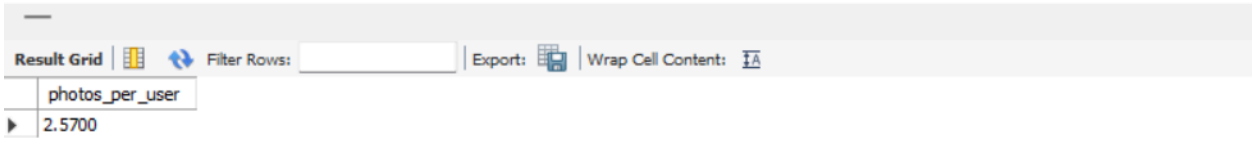
Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

registration_day	user_count
Thursday	16

6. User Engagement Metrics: The average number of posts per user and the total number of photos per user ratio were calculated, showing healthy user activity levels on the platform.

```
126 •   SELECT AVG(photo_count) AS avg_posts_per_user
127     FROM (
128         SELECT COUNT(p.id) AS photo_count
129         FROM users u
130         LEFT JOIN photos p ON u.id = p.user_id
131         GROUP BY u.id
132     ) AS user_photo_counts;
133
```


```
134 •   SELECT (SELECT COUNT(*) FROM photos) / (SELECT COUNT(*) FROM users) AS photos_per_user;
135
136
```



7. Bot Detection: Potential bot accounts were identified by flagging users who liked every photo on the platform, helping maintain the integrity of the user base.

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

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

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

Conclusion-

This project demonstrated how SQL can be effectively used to extract valuable insights from Instagram's user data. By analyzing various aspects such as user loyalty, engagement, hashtag usage, and potential bot activity, I was able to provide actionable recommendations that can enhance Instagram's marketing strategies and platform optimization. The identification of inactive users and the optimal days for ad campaigns will help improve engagement, while detecting bots will contribute to maintaining a genuine user base. These insights can guide the product and marketing teams in making informed, data-driven decisions to foster growth and improve user experience on one of the world's leading social media platforms.