Instagram User Analytics

Presented by: Suhita Vaidya

Project description

The aim of this project is to employ User analysis to track how users engage and interact with our digital product (software or mobile application) in an attempt to derive business insights for marketing, product & development teams.

This project focuses on two aspects:

- Marketing metrics:

–Rewarding Most Loyal Users
–Remind Inactive Users to Start Posting
–Declaring Contest Winner
–Hashtag Researching
–Launch AD Campaign

- Investor metrics:

–User Engagement
–Bots & Fake Accounts

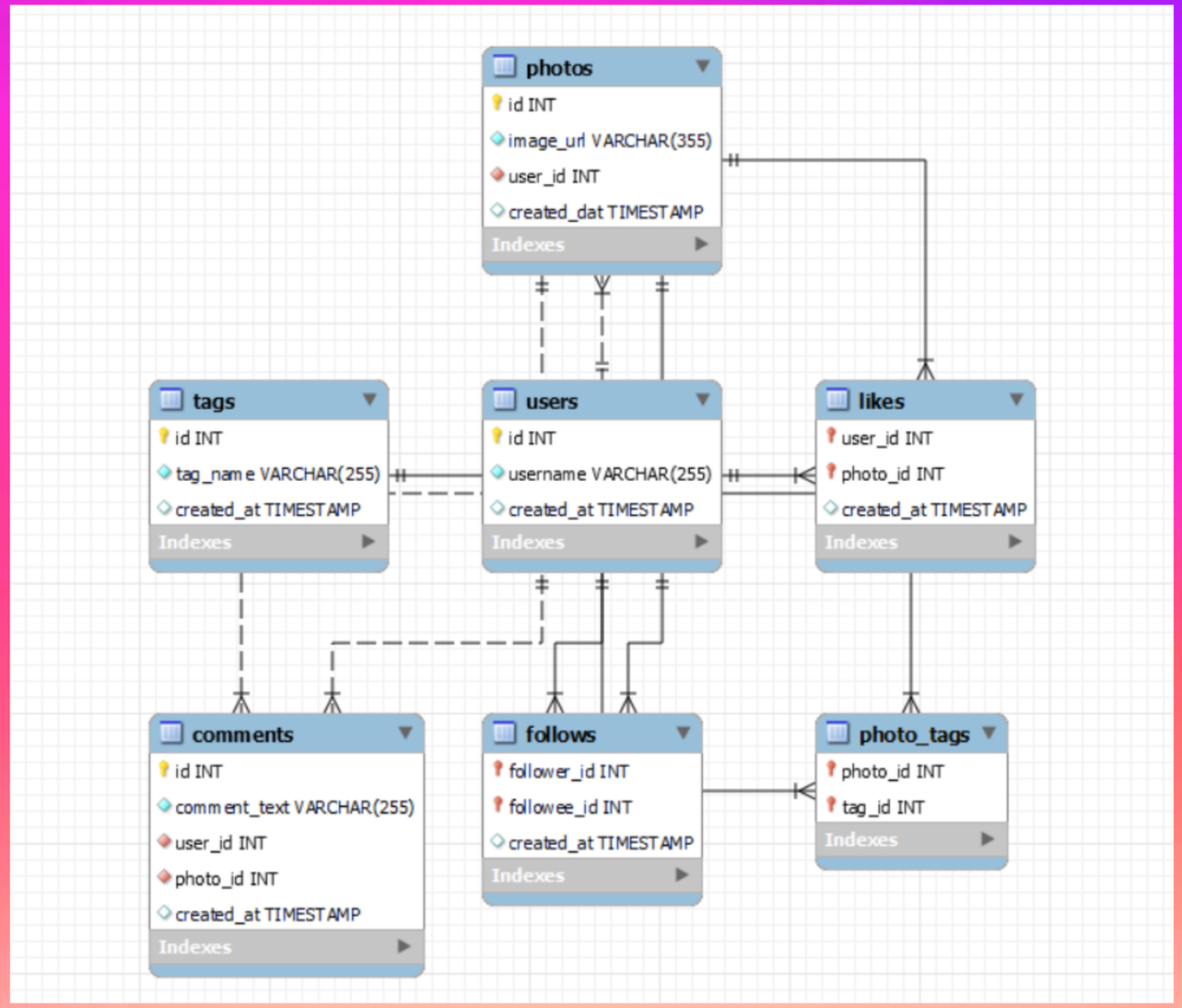
Approach:

First, the database is created in MySQL Workbench by executing queries. The datasets were studied with the help of ER diagram to be used for analysis.

Tech stack used:

MySQL Workbench was used to create the database and store all the records. It is also used to perform the required analysis using SQL queries.

ER Diagram


photos
id INT
image_url VARCHAR(355)
user_id INT
created_dat TIMESTAMP
Indexes
tags
id INT
tag_name VARCHAR(255)
created_at TIMESTAMP
Indexes
users
id INT
username VARCHAR(255)
created_at TIMESTAMP
Indexes
likes
user_id INT
photo_id INT
created_at TIMESTAMP
Indexes
comments
id INT
comment_text VARCHAR(255)
user_id INT
photo_id INT
created_at TIMESTAMP
Indexes
follows
follower_id INT
followee_id INT
created_at TIMESTAMP
Indexes
photo_tags
photo_id INT
tag_id INT
Indexes

MARKETING METRICS

The marketing team wants to launch some campaigns and they need some help with the following metrics.

Insights

1) Rewarding Most Loyal Users: People who have been using the platform for the longest time.

Task: Find the 5 oldest users of Instagram from the database provided.

➤ Here are the top 5 most loyal users with their user IDs and registration dates.

```
1 •   SELECT
2         *
3     FROM
4         users
5     ORDER BY created_at DESC
6     LIMIT 5;
```

id	username	created_at
11	Justina.Gaylord27	2017-05-04 16:32:16
6	Travon.Waters	2017-04-30 13:26:14
85	Milford_Gleichner42	2017-04-30 07:50:51
19	Hailee26	2017-04-29 18:53:40
24	Maxwell.Halvorson	2017-04-18 02:32:44

2) Remind Inactive Users to Start Posting:
By sending them promotional emails to post their 1st photo.

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

➤ There are 100 users out of which only 74 have posted pictures and 26 have never posted a single picture on Instagram.

```
1  SELECT
2      e.id, e.username
3  FROM
4      users e
5  WHERE
6      e.id NOT IN (SELECT
7              p.user_id
8          FROM
9              photos p)
```

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

3) Declaring Contest Winner :

The team started a contest and the user who gets the most likes on a single photo will win the contest. Now they wish to declare the winner.

Task: Identify the winner of the contest and provide their details to the team

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

➤ The contest winner is Zack_Kemmer93 with user ID 52 who got 48 likes on a single photo.

	id	username	photo_id	Most_likes
▶	52	Zack_Kemmer93	145	48

4) Hashtag Researching :

A partner brand wants to know which hashtags to use in a post to reach most people on the platform

Task: Identify and suggest the top 5 most commonly used hashtags on the platform

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

tag_id	tag_name	No_of_tags
21	smile	59
20	beach	42
17	party	39
13	fun	38
18	concert	24
5	food	24
11	lol	24

➤These are the 5 most commonly used hashtags on the platform–
smile, beach, party, fun, concert, food, lol.

5) Launch AD Campaign :

The team wants to know which day would be the best day to launch the ADs.

Task: What day of the week do most users register on? Provide insights on when to schedule an ad campaign

```sql
SELECT
    DAYNAME(created_at) AS Day_name, COUNT(id) AS Users
FROM
    users
GROUP BY Day_name
ORDER BY Users DESC;
```

Day_name	Users
Thursday	16
Sunday	16
Friday	15
Tuesday	14
Monday	14
Wednesday	13
Saturday	12

➤Thursdays and Sundays were the days when most users registered

INVESTOR METRICS

Our investors want to know if Instagram is performing well and is not becoming redundant like Facebook, they want to assess the application on the following grounds

1) User Engagement: Are users still as active and post on Instagram or they are making fewer posts?

Task: Provide how many times does average user posts on Instagram. Also, provide the total number of photos on Instagram/total number of users

➤ Average posts per user is around 3.47 with the total user count as 100 and post count as 257

```sql
SELECT
    AVG(count) AS Average_posts_per_user
FROM
    (SELECT
        user_id, COUNT(id) AS count
    FROM
        photos
    GROUP BY user_id) AS temp;
```

Average_user_posts
3.4730

```sql
SELECT
    COUNT(id) AS User_count
FROM
    users;
```

User_count
100

```sql
SELECT
    COUNT(id) AS Post_count
FROM
    photos;
```

Post_count
257

2) Bots & Fake Accounts: The investors want to know if the platform is crowded with fake and dummy accounts

Task: Provide data on users (bots) who have liked every single photo on the site (since any normal user would not be able to do this).

```sql
SELECT
    id, username AS Bots_name, COUNT(photo_id) AS Likes_on_posts
FROM
    likes
        JOIN
    users ON users.id = likes.user_id
GROUP BY likes.user_id
HAVING Likes_on_posts = MAX(photo_id);
```

id	Bots_name	Likes_on_posts
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	Leslie67	257
76	Janelle.Nikolaus81	257
91	Bethany20	257

➤ There are 13 users who have liked every post on the platform since their likes count is equal to the total post count.

RESULTS

- Proficiency in SQL: Developed a strong command over SQL queries, including advanced techniques for data retrieval, filtering, and aggregation.
- Data exploration: Navigated complex Instagram databases using SQL to extract relevant information for marketing campaigns and investor metrics.
- User segmentation: Utilized SQL queries to categorize users based on their activity levels, such as identifying the oldest users and inactive users.
- Performance analysis: Analyzed user engagement metrics, such as average post frequency, to assess the level of user activity on Instagram.
- Contest management: Leveraged SQL to identify the contest winner by analyzing the number of likes on individual photos.
- Hashtag research: Utilized SQL queries to identify the most commonly used hashtags on the platform, providing insights for effective marketing strategies.
- Timing optimization: Utilized SQL queries to identify the day of the week with the highest user registrations, helping determine the best day to launch advertising campaigns.
- Identifying bots and fake accounts: Employed SQL queries to identify users who exhibited suspicious behavior, such as liking every single photo on the site, indicating the presence of potential bots.
- Business decision-making- I learned about how few insights can have a huge impact on the business decision-making process and how real-life business case scenarios are.

Thank you.