INSTAGRAM USER ANALYTICS

PROJECT DESCRIPTION

Imagine you're a data analyst working with the product team at Instagram. Your role involves analyzing user interactions and engagement with the Instagram app to provide valuable insights that can help the business grow. In this project, you'll be using SQL and MySQL Workbench as your tool to analyze Instagram user data and answer questions posed by the management team. Your insights will help the product manager and the rest of the team make informed decisions about the future direction of the Instagram app. Remember, the goal of this project is to use your SQL skills to extract meaningful insights from the data.

APPROACH

I started by learning about relational databases, SQL commands and how to query a database to get meaningful information. I practiced and solved questions based on these commands utilizing online resources such as sqlzoo and sql-practice.

Next, I explored the given database file and followed the commands to create the tables and insert all the values. Then, I drew the Database schema diagram online on drawsql.app.

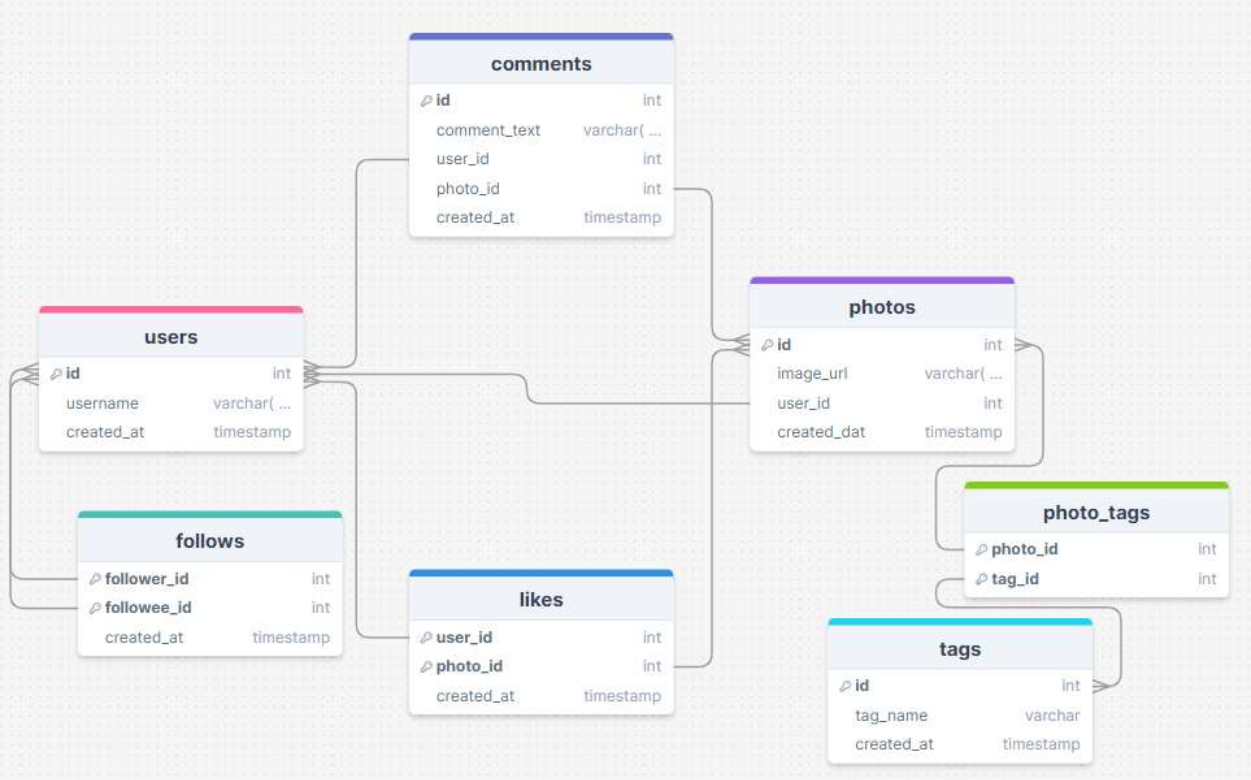
Once the schema was ready, I analysed the tables, their data and references such as PRIMARY KEYS and FOREIGN KEYS. Next, I used the MySQL Workbench to implement the queries using Join function, subqueries, Aggregation, where condition, Group by, Distinct and other functions to extract the required data from the database ensuring accuracy and efficiency. The results provided with meaningful insights to present to the team.

TECH-STACK USED

- ✓ Operating System – Microsoft Windows 11
- ✓ MySQL Workbench – To create database and write queries
- ✓ MySQL Server
- ✓ MySQL Shell
- ✓ Drawsql.app – To draw the schema diagram
- ✓ Microsoft Word – To write Report

INSIGHTS

Database Schema Diagram:



A) Marketing Analysis:

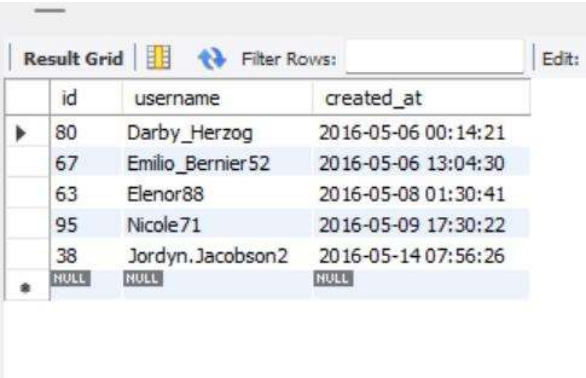
1) LOYAL USER REWARD

The marketing team wants to reward the most loyal users, i.e., those who have been using the platform for the longest time.

SQL Query:

```
# 1) Identify the five oldest users on Instagram from the provided database.
select * from users
order by created_at asc
limit 5;
```

Result:



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) <u>INACTIVE USER ENGAGEMENT</u>

The team wants to encourage inactive users to start posting by sending them promotional emails.

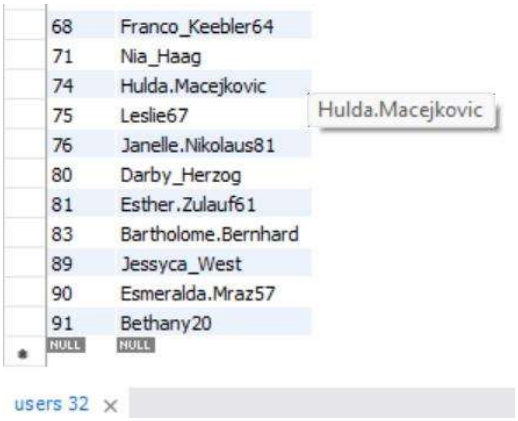
SQL Query:

```
# 2) Identify users who have never posted a single photo on Instagram
select id, username from users
where id not in (select user_id from photos);
```

Result:

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

users 32



3) CONTEST WINNER DECLARATION

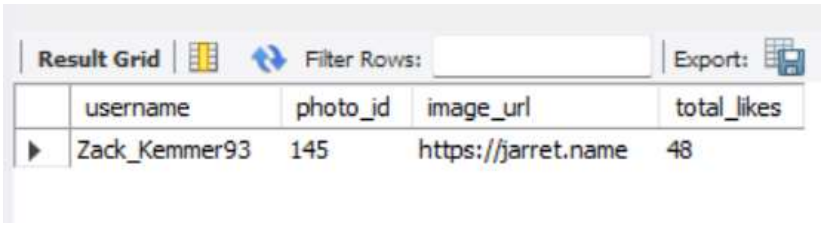
The team has organized a contest where the user with the most likes on a single photo wins.

SQL Query:

```
# 3) Determine the winner of the contest and provide their details to the team.
# The user with the most likes on a single photo wins.

select username, photos.id as photo_id , photos.image_url, count(likes.user_id) 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;
```

Result:



4) HASHTAG RESEARCH

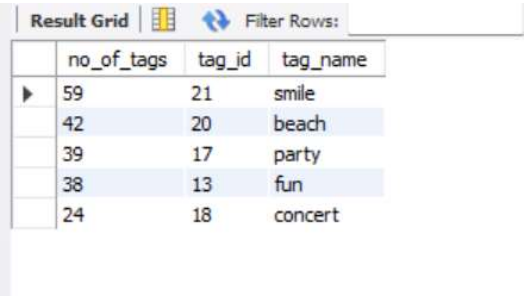
A partner brand wants to know the most popular hashtags to use in their posts to reach the most people.

SQL Query:

```
# 4) Identify and suggest the top five most commonly used hashtags on the platform.

select
   count(photo_id) as no_of_tags,
   tag_id,
   tag_name
from
   photo_tags p join tags t
   on p.tag_id = t.id
group by tag_id
order by no_of_tags desc
limit 5;
```

Result:



Result Grid | Filter Rows:

no_of_tags	tag_id	tag_name
59	21	smile
42	20	beach
39	17	party
38	13	fun
24	18	concert

5) AD CAMPAIGN LAUNCH

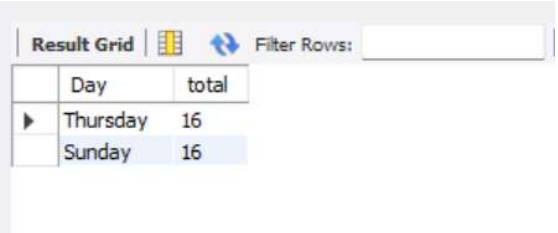
The team wants to know the best day of the week to launch ads.

SQL Query:

```
# 5) Determine the day of the week when most users register on Instagram. Provide insights on when to schedule an ad campaign.

select dayname(created_at) as Day, count(*) as total
from users
group by Day
order by total desc
limit 3;
```

Result:



Result Grid | Filter Rows:

	Day	total
▶	Thursday	16
	Sunday	16

B) <u>Investor Metrics:</u>

1) <u>USER ENGAGEMENT:</u>

Investors want to know if users are still active and posting on Instagram or if they are making fewer posts.

SQL Query:

```
# 6) Calculate the average number of posts per user on Instagram. Also, provide the total number of photos on Instagram divided

select
    avg(num_of_post) as avg_post_per_user
from
    (select
        user_id, COUNT(*) as num_of_post
    from
        photos
    group by user_id) as average_post_count;

select
    (select count(*) from photos) /(select count(*) from users)
as Total_photos_by_total_users;
```

Result:

Result Grid | Filter

	avg_post_per_user
▶	3.4730

Result Grid | Filter Rows:

	Total_photos_by_total_users
▶	2.5700

2) BOTS & FAKE ACCOUNTS

Investors want to know if the platform is crowded with fake and dummy accounts.

SQL Query:

```
# 7)Identify users (potential bots) who have liked every single photo on the site,

select user_id, username
from likes join users
on likes.user_id = users.id
group by user_id
having count(photo_id) = (select count(*) from photos);
```

Result:

Result Grid | Filter Rows:

	user_id	username
▶	5	Aniya_Hackett
	14	Jadyn81
	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

RESULT

The insights I found from this project are as follows:

- ✓ The top 5 oldest users
- ✓ To encourage those users who never posted photos through emails and contests
- ✓ Most liked photo who is declared as winner of the contest and conduct more such events to increase user level interaction
- ✓ Top 5 Hash tags to be used by users to reach and gain followers by increasing their visibility
- ✓ Best day to schedule an ad campaign to reach more users
- ✓ Total number of users
- ✓ Total number of photos
- ✓ Average number of photos per user
- ✓ Bots and fake accounts

From this project, I got a real time experience of working as a data analyst. I have gained a lot of understanding and knowledge on fundamentals of SQL.

This case study helped me to further test my knowledge on SQL and work on real life example queries.

It helped me demonstrate my critical thinking and gain better understanding of analysis process by answering the case study questions. Overall, it was a good experience to do hands on project.