INSTAGRAM User Analytics

Deovart Verma | Project 2

A) **PROJECT DESCRIPTION:**
This project is done as per the requirements specified; the major aim of this project is to generate insights from the users' database of Instagram using query management tool MySQL.

B) **APPROACH:**
For execution of this project, we used MySQL workbench, then created a database from the provided doc. Then, we analysed all the tables and schema in the database and the details in them. Then executed SQL queries as per the requirements posted.

C) **TECH-STACK USED:**
For the completion of this project we used MySQL workbench 8.0 CE, it a user-friendly tool to create SQL databases and execute SQL queries on it.

Q.A) MARKETING ANALYSIS:

#1 Identify the five oldest users on Instagram from the provided database.

```
select username, created_at from users order by created_at ASC LIMIT 5;
```

Insight: These are the oldest users of the platform.

username	created_at
Darby_Herzog	2016-05-06 00:14:21
Emilio_Bernier52	2016-05-06 13:04:30
Elenor88	2016-05-08 01:30:41
Nicole71	2016-05-09 17:30:22
Jordyn.Jacobson2	2016-05-14 07:56:26

#2 Identify users who have never posted a single photo on Instagram.

```
select id, username from users where id not in (select distinct(user_id) from photos);
```

Insight: These are the platform users that haven't posted yet on the platform, these can be targeted for promotional strategies to make them active on the platform.

id	username
5	Aniya_Hackett
7	Kasandra_Homenick
14	Jaclyn81
21	Rocio33
24	Maxwell.Halvorson
25	Tierra.Trantow

#3 Determine the winner of the contest and provide their details to the team.

```
SELECT username, photos.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;

Insight: This user has most number of likes in a single post on the entire platform.

username	id	image_url	total_likes
Zack_Kemmer93	145	https://jarret.name	48

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

SELECT tags.tag_name, Count(*) AS tag_count

FROM photo_tags

JOIN tags

ON photo_tags.tag_id = tags.id

GROUP BY tags.id

ORDER BY tag_count DESC LIMIT 3;

Insight: Out of all the tags in the platform, these are the tags that were used the most.

tag_name	tag_count
smile	59
beach	42
party	39

#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 2;

Insight: According to the insights generated these are the days, on which maximum users have registered on the platform, so the team can schedule promotional strategies on these days.

day	total
Thursday	16
Sunday	16

#B. Investors Metrics

#1 Calculate the average number of posts per user on Instagram. Also, provide the total number of photos on Instagram divided by the total number of users.

SELECT (SELECT COUNT(*) FROM photos) / (SELECT COUNT(*) FROM users) AS avg;

Insight: Looking at the average number of posts on the platform, we can that the users on the platform are active and are enjoying the platform.

avg
2.5700

#2 Identify users (potential bots) who have liked every single photo on the site, as this is not typically possible for a normal user.

SELECT users.username, COUNT(*) as num_likes

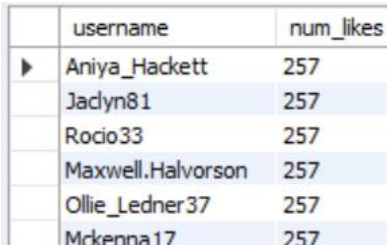
FROM users

JOIN likes ON users.id = likes.user_id

GROUP BY users.id

HAVING num_likes = (SELECT COUNT(*) FROM photos);

Insight: These account seem like bot accounts, as they have liked entire photos on the platform.



username	num_likes
Aniya_Hackett	257
Jadyn81	257
Rocio33	257
Maxwell.Halvorson	257
Ollie_Ledner37	257
Mckenna17	257

<u>RESULT:</u> After the completion of the project, my grasp on the SQL queries have improved, and I've got an idea how SQL is used in real world for generating insights.