Instagram user Analytics

SQL Fundamentals Project Report

Roopa M

1.Project Overview

Objective:

The goal of this project was to understand how Instagram users interact with the app. By exploring the data using SQL, I wanted to uncover useful insights that could help marketing teams, product managers, and investors make informed decisions.

Description:

I worked with a dataset containing user profiles, posts, likes, and hashtags. The aim was to figure out which users are the most loyal, who hasn't posted yet, which posts get the most attention, the most popular hashtags, and general trends in user engagement. These insights are valuable for boosting campaigns, improving engagement, and spotting potential bot accounts.

2. Approach

1. set up the Instagram database in MySQL workbench using SQL file provided.
2. Explored each table carefully: users, comments, follows, photos, likes, tags, and photo_tags .
3. Wrote SQL queries to answer each business question step by step.
4. Collected the query results and interpreted what they meant for the business.
5. Compiled the queries, results, and insights into this report for clarity.

3. Tech-Stack Used

- **MySQL Workbench 8.0**- to run queries and check outputs.
- **MySQL Server**- to store and manage the database.
- **Ms Word/PDF**- for organising and presenting the report.

Why I choose these tools:

- MySQL Workbench is user-friendly and makes quering simple.
- Word/pdf keeps everything clean, organised, and easy to share.

4. Analysis and Findings

A) Marketing Insights

1.Identifying Loyal Users- Oldest 5 Accounts

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

Result

id	username	created_at
80	Darby_Herzog	06-05-2016 0:14
67	Emilio_Bernier52	06-05-2016 13:04
63	Elenor88	08-05-2016 1:30
95	Nicole71	09-05-2016 17:30
38	Jordyn.Jacobson2	14-05-2016 7:56

Insight:

These users joined first and have been active for a long time. Offering them loyalty perks could encourage them to stay even more engaged.

2. engaging inactive users- Those who never posted

```
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
71	Nia_Haag
74	Hulda.Macejkovic
75	Leslie67
76	Janelle.Nikolaus81
80	Darby_Herzog
81	Esther.Zulauf61
83	Bartholome.Bernhard
89	Jessyca_West
90	Esmeralda.Mraz57
91	Bethany20

Insight:

26 users have not posted any photos.

3. contest winner declaration -most likes on a single photo

```
SELECT id, username, created_at
FROM users
WHERE id = (SELECT user_id
FROM photos
WHERE id = (SELECT photo_id
FROM likes
GROUP BY photo_id
ORDER BY COUNT (photo_id ) DESC
LIMIT 1)
);
```

Result:

id	username	created_at
52	Zack_Kemmer93	01-01-2017 5:58

Insights:

The user whose photo received the maximum likes was declared the winner.

4. Most popular hashtags research

```
SELECT tags. tag_name,
COUNT (tags.tag_name) AS "tags_count"
FROM tags
INNER JOIN photo_tags
ON tags.id = photo_tags.tag_id
GROUP BY tags.tag_name
ORDER BY COUNT (tags.tag_name) DESC
LIMIT 5;
```

Result:

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

Insight:

These hashtags are currently trending.

5. Best days for Ad campaigns

```
SELECT DAYNAME (created_at) AS registration_day,
COUNT (DAYNAME (created_at) )  AS `Count of register`
FROM users
GROUP BY DAYNAME (created_at)
ORDER BY COUNT (DAYNAME (created_at)) DESC
LIMIT 7;
```

Result:

registration_day	Count of register
Thursday	16
Sunday	16
Friday	15
Tuesday	14
Monday	14
Wednesday	13
Saturday	12

Insight:

By analysing user registration dates, I found which day of week had the most new registrations. That day can be considered the best for launching promotional ads or marketing events.

B) Investor Metrics

1. User Engagement

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

Result

avg
2.5700

Insights

I calculated the average number of posts per user. This metric helps measure how active users are on Instagram.

Formula used: (total number of photos / total number of users)

2. Bots and fake Accounts

```
SELECT id, username, created_at
FROM users
WHERE id IN (SELECT user_id
FROM likes
GROUP BY user_id
HAVING COUNT (user_id ) = (SELECT COUNT (id)
FROM photos)
);
```

Result

id	username	created_at
5	Aniya_Hackett	07-12-2016 1:04
14	Jaclyn81	06-02-2017 23:29
21	Rocio33	23-01-2017 11:51
24	Maxwell.Halvorson	18-04-2017 2:32
36	Ollie_Ledner37	04-08-2016 15:42
41	Mckenna17	17-07-2016 17:25
54	Duane60	21-12-2016 4:43
57	Julien_Schmidt	02-02-2017 23:12
66	Mike.Auer39	01-07-2016 17:36

71	Nia_Haag	14-05-2016 15:38
75	Leslie67	21-09-2016 5:14
76	Janelle.Nikolaus81	21-07-2016 9:26
91	Bethany20	03-06-2016 23:31

Insight

These are likely bot accounts. Removing them ensures the platform's engagement data is accurate.

5. Summary of Key Insights

Area	What I found	What I suggest
Loyal users	Found the 5 oldest users on instagram .	Give them a reward to keep them active and loyal.
Inactive users	26 users have never posted anything.	Send them a remainder or email to get them started
Contest winner	One user got the most likes on a single post.	Highlight them or feature their post to inspire others.
Hashtags	Most used hashtags	Use these in marketing or brand posts.
Ad timing	Most signups happen on Thursday and Sunday	Run ads on those days for better reach
Engagement	On average, user post around 2.5 photos	Try to get users to post more often
Fake accounts	13 users liked every single photo.	Check and remove them to keep data clean.

6. Result

- Practiced SQL skills including **JOINS, GROUP BY, CTE's, and subqueries**.
- Learned how to turn raw data into insights that can guide business decisions.
- Gained experience in interpreting engagement trends and spotting irregular activity.
- Built analytical reasoning and business perspective in real-world scenarios.

7. Drive link

Google drive link:
https://drive.google.com/file/d/1yqQfnS_oHstMPt04ojTxYaDmVrlNZqJN/view?usp=sharing

8. Conclusion

This project shows how SQL can provide actionable insights from Instagram user data. I identified loyal users, engagement, patterns, trending hashtags, and potential bot accounts. These findings can help marketing strategies, improve user engagement, and provide reliable metrics for investors.