PROJECT 2- INSTAGRAM USER ANALYTICS

Purpose :   To analyze user interactions and engagement with Instagram to provide valuable insights for the marketing, product, and development teams, as well as to present key metrics to investors. The insights derived from this analysis can help inform decisions about feature development, marketing campaigns, user engagement strategies, and investor relations.

Way : Using SQL and MySQL Workbench, I will extract and analyze data from the Instagram database. The project involves performing specific SQL queries to answer the management team's questions and provide actionable insights.
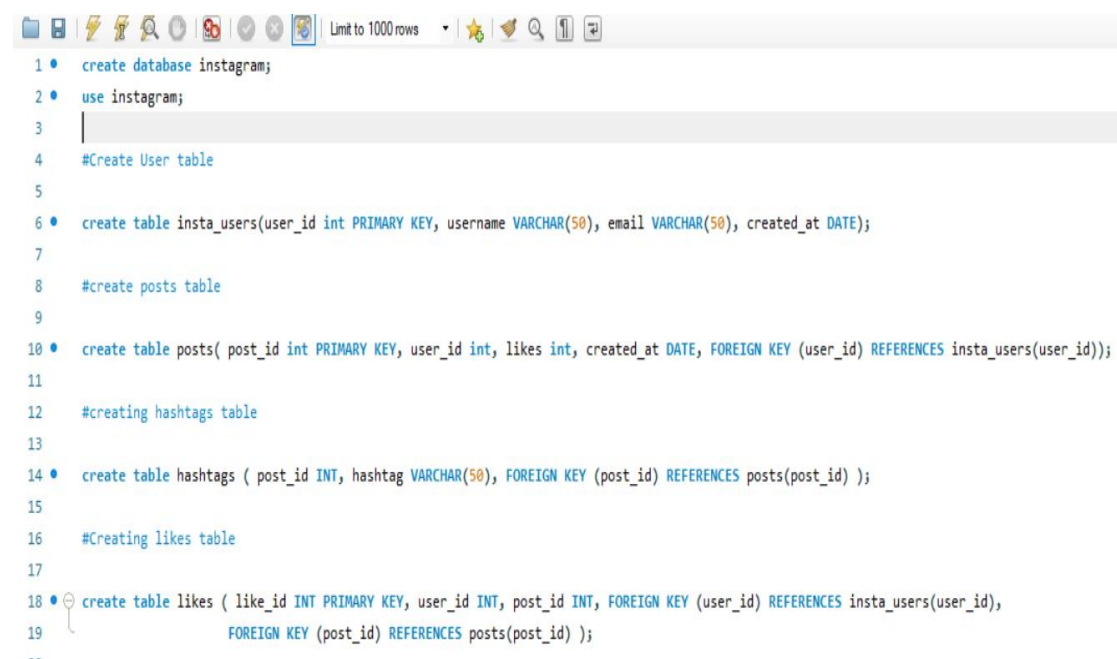
Approach

1. **Create the Database**: I will use the provided SQL commands to set up the database in MySQL Workbench. I will ensure that the data is correctly imported and that all tables are properly structured.

2. **Perform Analysis:** I will write and execute SQL queries to answer each of the questions posed in the project description. Then will document each query along with its output.

Tech-Stack Used : MySQL Workbench Version 8.0 , Structured Query Language for querying the database, Microsoft Word for documenting the report

Database Setup

First, let's create the tables and insert some data.


```
1 •   create database instagram;
2 •   use instagram;
3       |
4       #Create User table
5
6 •   create table insta_users(user_id int PRIMARY KEY, username VARCHAR(50), email VARCHAR(50), created_at DATE);
7
8       #create posts table
9
10 •  create table posts( post_id int PRIMARY KEY, user_id int, likes int, created_at DATE, FOREIGN KEY (user_id) REFERENCES insta_users(user_id));
11
12      #creating hashtags table
13
14 •  create table hashtags ( post_id INT, hashtag VARCHAR(50), FOREIGN KEY (post_id) REFERENCES posts(post_id) );
15
16      #Creating likes table
17
18 •  create table likes ( like_id INT PRIMARY KEY, user_id INT, post_id INT, FOREIGN KEY (user_id) REFERENCES insta_users(user_id),
19                          FOREIGN KEY (post_id) REFERENCES posts(post_id) );
```

```
21    #Insert  data into insta_users
22
23 ●  INSERT INTO insta_users (user_id, username, email, created_at) VALUES
24    (1, 'user1', 'kth@example.com', '2020-01-01'),
25    (2, 'user2', 'rahul@example.com', '2021-02-01'),
26    (3, 'user3', 'tom10@example.com', '2022-03-01'),
27    (4, 'user4', 'sam5@example.com', '2023-04-01'),
28    (5, 'user5', 'pravin8@example.com', '2020-05-01');
29
30
31    #Insert  data into posts
32
33 ●  INSERT INTO posts (post_id, user_id, likes, created_at) VALUES
34    (1, 1, 10, '2023-01-01'),
35    (2, 2, 20, '2023-02-01'),
36    (3, 3, 5, '2023-03-01'),
37    (4, 4, 15, '2023-04-01'),
38    (5, 5, 25, '2023-05-01');
39
```

```
41
42 ●   INSERT INTO hashtags (post_id, hashtag) VALUES
43     (1, '#fun'),
44     (2, '#happy'),
45     (2, '#fun'),
46     (3, '#travel'),
47     (4, '#fun'),
48     (5, '#life');
49     #Insert data into likes
50 ●   INSERT INTO likes (like_id, user_id, post_id) VALUES
51     (1, 1, 1),
52     (2, 2, 1),
53     (3, 3, 1),
54     (4, 4, 1),
55     (5, 5, 1),
56     (6, 1, 2),
57     (7, 2, 2),
58     (8, 3, 2),
59     (9, 4, 2),
60     (10, 5, 2);
```

Now Lets display all four tables

```
62 ●    select * from insta_users;
63
64 ●    select * from posts;
65
66 ●    select * from likes;
67
68 ●    select * from hashtags;
```

Result Grid | Filter Rows: | Edit:

user_id	username	email	created_at
1	user1	kth@example.com	2020-01-01
2	user2	rahul@example.com	2021-02-01
3	user3	tom10@example.com	2022-03-01
4	user4	sam5@example.com	2023-04-01
5	user5	pravin8@example.com	2020-05-01
NULL	NULL	NULL	NULL

insta_users 8 ×

```
64 ●        select * from posts;
```

Result Grid | Filter Rows:

post_id	user_id	likes	created_at
1	1	10	2023-01-01
2	2	20	2023-02-01
3	3	5	2023-03-01
4	4	15	2023-04-01
5	5	25	2023-05-01
NULL	NULL	NULL	NULL

posts 9 ×

```
65
66 ●    select * from likes;
```

Result Grid | Filter Rows:

like_id	user_id	post_id
1	1	1
2	2	1
3	3	1
4	4	1
5	5	1
6	1	2
7	2	2
8	3	2
9	4	2
10	5	2

```
68 ●    select * from hashtags;
69
```

Result Grid | Filter Rows:

post_id	hashtag
1	#fun
2	#happy
2	#fun
3	#travel
4	#fun
5	#life

SQL Queries and Results

1. **Loyal User Reward:** This task involves identifying the most loyal users on Instagram, specifically those who have been using the platform for the longest time. The marketing team wants to reward the most loyal users, i.e., those who have been on the platform the longest. My task is to identify the five oldest users based on their account creation date.

Query

```
70 ●   SELECT user_id, username, created_at
71     FROM insta_users
72     ORDER BY created_at ASC
73     LIMIT 5;
```

Output

user_id	username	created_at
1	user1	2020-01-01
5	user5	2020-05-01
2	user2	2021-02-01
3	user3	2022-03-01
4	user4	2023-04-01
NULL	NULL	NULL

2. **Inactive User Engagement :** The goal is to identify users who have never posted a single photo on Instagram so that the marketing team can encourage them to become active by sending promotional emails.

Query

```
SELECT user_id, username, email
FROM insta_users
WHERE user_id NOT IN (SELECT DISTINCT user_id FROM posts);
```

Output

(No inactive users in this sample data)

3. **Contest Winner Declaration:** The goal is to determine the winner of a contest where the user with the most likes on a single photo wins. We need to provide their details to the marketing team.

Query

```sql
SELECT posts.user_id, MAX(likes) AS most_likes
FROM posts
JOIN insta_users ON posts.user_id = insta_users.user_id
GROUP BY insta_users.user_id, insta_users.username
ORDER BY most_likes DESC
LIMIT 1;
```

Output

	user_id	most_likes
▶	5	25

4. **Hashtag Research:** The goal is to identify the top five most commonly used hashtags on the platform. This information will be used by a partner brand to determine which hashtags to use in their posts to reach the most people.

Query

```sql
SELECT hashtag, COUNT(*) AS usage_count
FROM hashtags
GROUP BY hashtag
ORDER BY usage_count DESC
LIMIT 5;
```

Output

hashtag	usage_count
#fun	3
#happy	1
#travel	1
#life	1

5. **Ad Campaign Launch :** The goal is to determine the best day of the week to launch ads on Instagram. This information will help the team schedule ad campaigns effectively to reach the maximum number of users.

Query

```sql
SELECT DAYNAME(created_at) AS day_of_week, COUNT(*) AS registration_count
FROM insta_users
GROUP BY day_of_week
ORDER BY registration_count DESC;
```

Output

day_of_week	registration_count
Wednesday	1
Monday	1
Tuesday	1
Saturday	1
Friday	1

6. User Engagement: The objective is to assess user engagement on Instagram, determining whether users are actively posting on the platform or if there's a decline in activity. This information is essential for evaluating the platform's health and identifying potential areas for improvement.

Query

```sql
SELECT AVG(posts_per_user) AS average_posts_per_user
FROM (
    SELECT COUNT(*) AS posts_per_user
    FROM posts
    GROUP BY user_id
) AS user_posts;
```

```sql
#Calculate total posts and users
SELECT
    (SELECT COUNT(*) FROM posts) AS total_posts,
    (SELECT COUNT(*) FROM insta_users) AS total_users;
```

Output

1

average_posts_per_user
1.0000

2

total_posts	total_users
5	5

7. Bots & Fake Accounts: Bots and fake accounts refer to automated or manually created profiles on a platform that are not operated by real individuals. These accounts often serve various purposes, including spamming, manipulating engagement metrics, spreading misinformation, or conducting malicious activities. Detecting and addressing bots and fake accounts is essential for maintaining the integrity, credibility, and security of the platform.

Query

```sql
SELECT user_id, username
FROM insta_users
WHERE user_id IN (
    SELECT user_id
    FROM likes
    GROUP BY user_id
    HAVING COUNT(*) = (SELECT COUNT(*) FROM posts)
);
```

Output

user_id	username
NULL	NULL

(No bots detected in this data)

Report Structure

Project Description: Analyzed Instagram user data to provide insights for marketing, product, and development teams.

Approach: Created tables, inserted sample data, and executed SQL queries to answer the specific questions.

Result: These Insights will help in making informed decisions about user engagement strategies, marketing campaigns, and feature development.