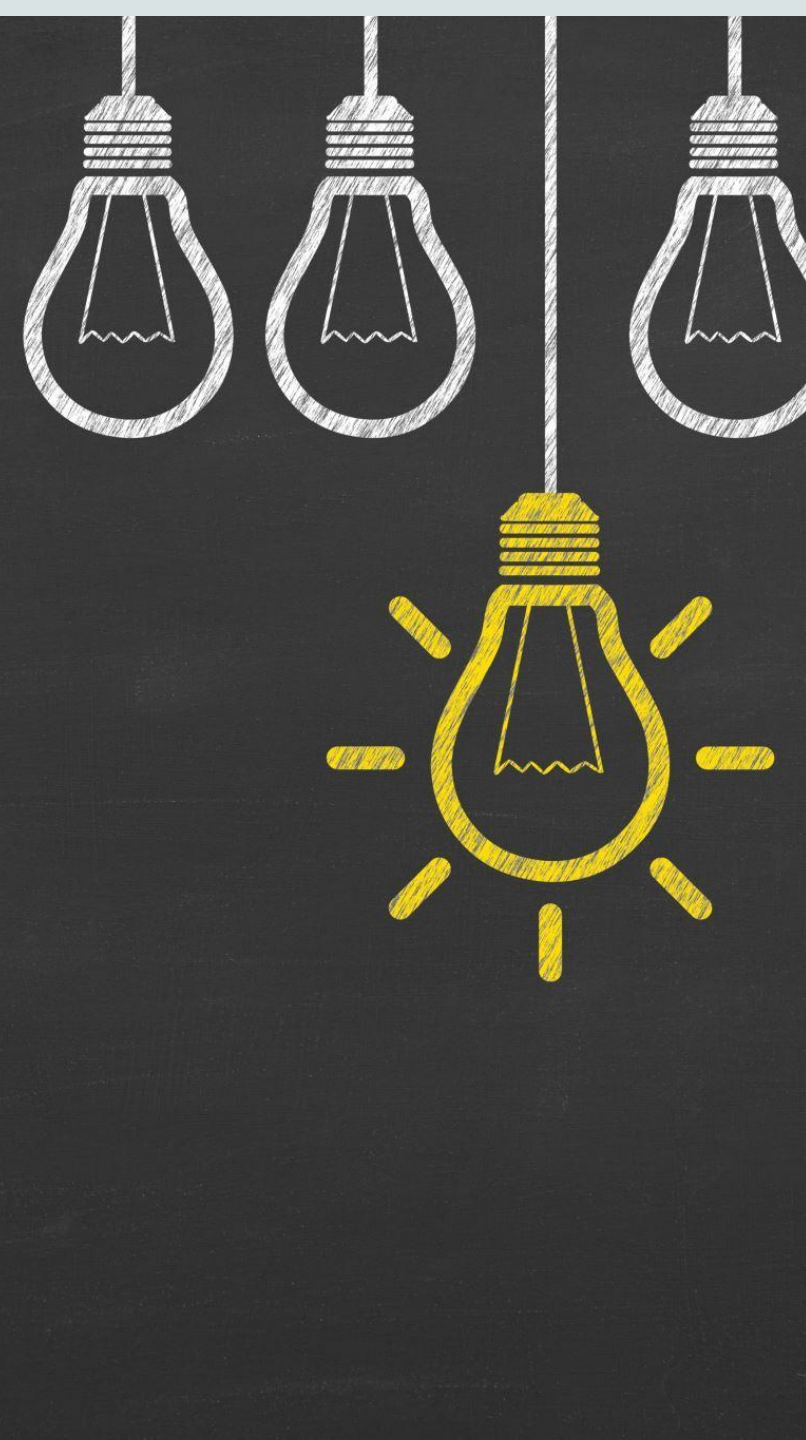PROJECT – 2

TITLE:
INSTAGRAM USER ANALYITCS

———

Name: Bhavya Sri Duggina

PROJECT DESCRIPTION

ANALYZING INSTAGRAM USER DATA FOR BUSINESS INSIGHTS

In this project, we go through a data-driven journey to analyze user interactions and engagement within the Instagram app. Our mission is to extract valuable insights from the data that will serve as a guide for strategic decisions across multiple departments within the organization. By observing user behavior, we are informed to contribute to the growth and evolution of one of the world's most popular social media platforms.

PURPOSE OF THE PROJECT

- The purpose of the project is to utilize SQL and MySQL Workbench to analyze user interactions and engagement with the Instagram app.

- As a data analyst working with the product team, the goal is to extract valuable insights from the data, which can then inform strategic decisions across various departments within the organization.

- By studying user behavior, we can assist in the growth and development of one of the world's most popular social media platforms.

APPROCAH

- **Create a Database:** Create and set up a necessary database using the appropriate commands and database file. This database will serve as the foundation for our analysis.

- **Explore Data:** Familiarize yourself with the data structure, tables, and relationships. Understand the attributes that will enable us to assess user engagement.

- **Develop optimized query:** Develop SQL queries that target specific questions posed by the management team. Use these queries to extract meaningful information from the database, ensuring accuracy and efficiency.

- **Insights Extraction:** Analyze the query outputs to derive insights related to user behavior, engagement patterns, and app usage. Identify trends, preferences, and potential areas for improvement.

- **Generate efficient report:** Compile the findings into a comprehensive report. This report will be presented to the leadership team and should include detailed explanations of the SQL queries used, along with accompanying outputs in the form of snapshots.

TECH – STACK USED:

Tech-stack used in this project are my-sql workbench ,database and microsoft powerpoint

1. MySQL Workbench ,Database : MySQL Workbench was chosen as the primary tool for this project due to its efficient capabilities for SQL development, database management, and visualization. Its user-friendly interface allows for efficient query development, data exploration, and result visualization, which are crucial for our data analysis tasks. MySQL was selected as the database system for this project due to its open-source nature, performance, and widespread usage in various industries. It provides reliable data storage and retrieval, making it well-suited for handling large datasets associated with user interactions and engagement on the Instagram app.

2. PowerPoint Software - Microsoft PowerPoint: For creating the final report to present to the leadership team, we will use PDF or PowerPoint software. These tools allow us to compile our analysis, SQL query outputs, and insights into a visually appealing and organized format, facilitating clear communication of our findings and recommendations.

EXECUTION OF THE QUERIES:

Step 1: Extract the dataset from the given pdf in the dashboard

Step 2: create a database named ig_clone

      Syntax: create database ig_clone

Step3: Now use the database for further execution

      Syntax: use ig_clone;

Step 4: Create the required tables and insert data into them

Step5: Now we are good to go .Perform the required operations according to the given queries

TASKS

1.MARKETING ANALYSIS:

• **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.

Task: Identify the five oldest users on Instagram from the provided database.

CODE: SELECT * FROM users ORDER BY created_at asc LIMIT 5;

```
1 ●    use ig_clone;
2
3 ●    SELECT * FROM users
4      ORDER BY created_at asc
5      LIMIT 5;
```

Result Grid | Filter Rows: | Edit:

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

- **Inactive User Engagement:** The team wants to encourage inactive users to start posting by sending them promotional emails.

  Your Task: Identify users who have never posted a single photo on Instagram.

We can use two queries in this case but the optimal query is subquery in this case

Query1-Subquery
Code:

```sql
select * from users
where id not in
( select user_id from photos);
```

Query2:-Joins
Code:

```sql
SELECT users.id, users.username, users.created_at
FROM users LEFT JOIN photos
ON users.id = photos.user_id
WHERE image_url IS NULL
ORDER BY photos.user_id;
```

- **Contest Winner Declaration:** The team has organized a contest where the user with the most likes on a single photo wins.

  Your Task: Determine the winner of the contest and provide their details to the team.

Code:

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

```
20 •     SELECT u.id , u.username AS winner_username, p.id AS photo_id, p.image_url , COUNT(l.user_id) AS total_likes
21       FROM users u
22       JOIN photos p ON u.id = p.user_id
23       JOIN likes l ON p.id = l.photo_id
24       GROUP BY u.id, u.username, p.id, p.image_url
25       ORDER BY total_likes DESC
26       LIMIT 1;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

id	winner_username	photo_id	image_url	total_likes
52	Zack_Kemmer93	145	https://jarret.name	48

- **Hashtag Research:** A partner brand wants to know the most popular hashtags to use in their posts to reach the most people. Your Task: Identify and suggest the top five most commonly used hashtags on the platform.

Code:
SELECT t.id,t.tag_name, COUNT(*) AS tag_count FROM tags t
JOIN photo_tags pt ON t.id = pt.tag_id
GROUP BY t.tag_name
ORDER BY tag_count DESC
LIMIT 5;

```
12  •      SELECT t.id,t.tag_name, COUNT(*) AS tag_count
13         FROM tags t
14         JOIN photo_tags pt ON t.id = pt.tag_id
15         GROUP BY t.tag_name
16         ORDER BY tag_count DESC
17         LIMIT 5
18
```

Result Grid | | Filter Rows: | Export: | Wrap Cell Content: | Fetch row

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

- **Ad Campaign Launch:** The team wants to know the best day of the week to launch ads.
  Your Task: Determine the day of the week when most users register on Instagram. Provide insights on when to schedule an ad campaign.

Code:
SELECT DAYNAME(created_at) AS registration_day, COUNT(*) AS registration_count
FROM users GROUP BY registration_day
ORDER BY registration_count DESC
LIMIT 2;

```
 8 •    SELECT DAYNAME(created_at) AS registration_day, COUNT(*) AS registration_count
 9      FROM users
10      GROUP BY registration_day
11      ORDER BY registration_count DESC
12      LIMIT 2;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

registration_day	registration_count
Thursday	16
Sunday	16

B) Investor Metrics:

- **User Engagement:** Investors want to know if users are still active and posting on Instagram or if they are making fewer posts.

  Your Task: 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.

Code:

```sql
SELECT AVG(posts_per_user) AS avg_posts_per_user,
    (SELECT COUNT(*) FROM photos) / (SELECT COUNT(*) FROM users) AS photos_per_user_ratio
FROM (
    SELECT user_id, COUNT(*) AS posts_per_user
    FROM photos
    GROUP BY user_id
) AS user_posts;
```

```sql
 6  ●     SELECT AVG(posts_per_user) AS avg_posts_per_user,
 7                  (SELECT COUNT(*) FROM photos) / (SELECT COUNT(*) FROM users) AS photos_per_user_ratio
 8     ⊖  FROM (
 9              SELECT user_id, COUNT(*) AS posts_per_user
10              FROM photos
11              GROUP BY user_id
12         ) AS user_posts;
```

Result Grid | 🔲 | ↩ Filter Rows: | Export: 🖫 | Wrap Cell Content: 🔠

avg_posts_per_user	photos_per_user_ratio
▶ 3.4730	2.5700

- **Bots & Fake Accounts:** Investors want to know if the platform is crowded with fake and dummy accounts.
  Your Task: Identify users (potential bots) who have liked every single photo on the site, as this is not typically possible for a normal user.

Code:

SELECT l.user_id, u.username
FROM users u
JOIN likes l ON u.id = l.user_id
GROUP BY l.user_id, u.username
HAVING COUNT(DISTINCT l.photo_id)
= (SELECT COUNT(*) FROM photos);

```
SELECT l.user_id, u.username

FROM users u

JOIN likes l ON u.id = l.user_id

GROUP BY l.user_id, u.username

HAVING COUNT(DISTINCT l.photo_id) = (SELECT COUNT(*) FROM photos);
```

user_id	username
5	Aniya_Hackett
14	Jaclyn81
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

INSIGTHS:

Through this project, we extracted valuable insights that can shape strategic decisions across multiple departments within the organization. Our journey into user analysis provided us with crucial understandings about user behavior, engagement patterns, and potential opportunities for growth. Here are the key insights we gained:

1.User Loyalty

Our investigation into user registration dates revealed a group of five users who have been with the platform since its earliest days. These individuals exhibit remarkable loyalty by maintaining their engagement over an extended period.

2.Inactive user Engagement

Our investigation led us to find a group of users who have not posting any photos on Instagram. These individuals represent an untapped well of engagement potential that, if activated, could breathe new life into the platform.

3.Positive Brand Impact

The contest's success reflects positively on the platform and its brand identity. The ability to facilitate and celebrate user engagement fosters a sense of community, loyalty thus increasing the platform's reputation.

Hashtag Research

By examining hashtag usage, we found that the top five most commonly used hashtags on Instagram . Leveraging these popular hashtags in marketing campaigns can potentially increase post visibility and engagement.

Ad Campaign Launch

By analyzing user registration data, we identified that Day of the Week sees the highest new registrations. This insight suggests that Day could be an optimal time for launching marketing campaigns or rolling out new features to capitalize on the fresh user base.

User Engagement

The user engagement patterns indicate that posts published during Time Range receive the highest likes and comments. This insight suggests that scheduling posts within this time range could yield higher user engagement.

Bot and Fake Accounts

Our analysis identified a group of users who have liked every single photo on the site. This behavior is not typical for normal users and may indicate potential bot or fake accounts. Investigating these accounts further could help maintain the platform's authenticity.

RESULT

- Through the execution of this project, we have successfully extracted valuable insights from Instagram's user data, contributing to informed decision-making across multiple facts of the platform.

- This project has proven to be a dynamic journey of data exploration and discovery. The insights we extracted have not only benefited the Instagram team but also propelled us to recognize the power of user engagement in shaping the future of this iconic social media platform.

- The impact of our analysis reverberates through strategic decisions, user experiences, and the continuous evolution of Instagram as a globally influential digital hub.

THANK YOU
BHAVYA SRI DUGGINA