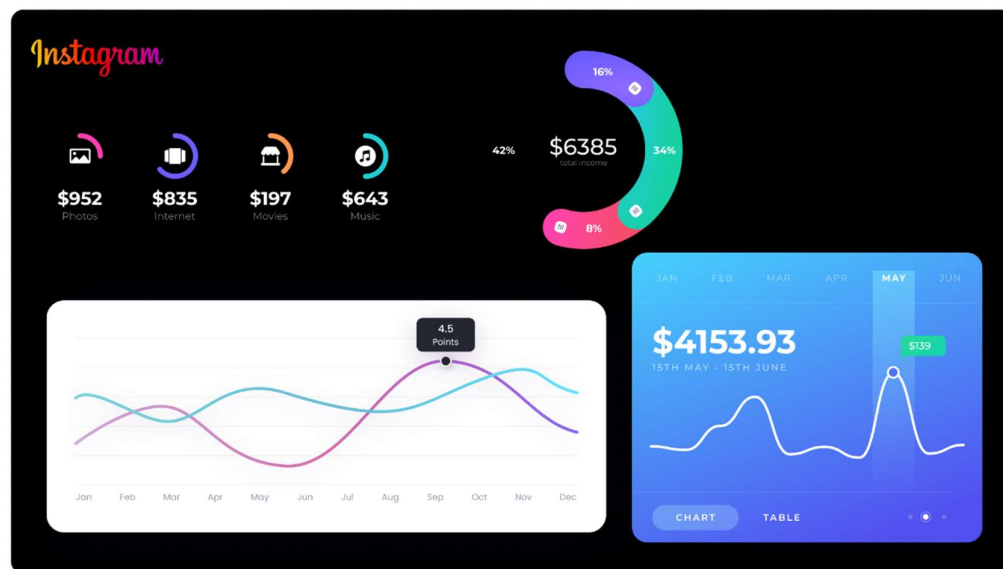# A
Live- Project Report
On
"Instagram User Analytics"

Presented to – Trainity

Submitted by-

Anirudh Chaudhary

Project Description:

This project focuses on analysing Instagram user interactions and engagement data to provide actionable insights that can drive platform growth. The analysis aims to support the marketing, product, and investor teams by uncovering trends and opportunities to enhance user experience and engagement.

Approach:

1. Exploring Data: The first step involved understanding the structure of the Instagram user database.

2. Using SQL: Queries were written to extract relevant insights based on specific project objectives.

3. Analyzing Results: The results were interpreted to generate actionable insights.

4. Presenting Findings: Insights were structured into professional slides for effective stakeholder communication.

Tools Used:

The project utilized **SQL** for data querying and **MySQL Workbench** as the database management tool to perform the analysis efficiently.

Approach:

The project follows a structured approach to analyse Instagram user data and extract meaningful insights   for the marketing, product, and investor teams. The steps are:

1. **Understanding the Objective:**

   ➢ The primary goal is to analyse user interactions to support business strategies.

   ➢ Specific tasks include identifying loyal users, understanding user activity trends, detecting bots, and optimizing marketing efforts.

2. **Data Exploration:**

   ➢ Examining the provided database schema to understand available tables, relationships, and attributes.

   ➢ Reviewing tables such as user profiles, posts, likes, and hashtags to identify relevant data points for each analysis task.

3. Query Execution Using SQL:

➢ Writing precise SQL queries to retrieve relevant data based on the problem statements.

➢ Ensuring that the queries are optimized for efficiency and accuracy.

4. Result Analysis:

➢ Interpreting the outputs of the queries to identify patterns, trends, and actionable insights.

➢ Cross-verifying results to ensure correctness and consistency.

5. Visualization and Presentation:

➢ Preparing screenshots of SQL query results for transparency.

➢ Structuring insights in a professional slide deck format for effective communication with stakeholders.

A) <u>Marketing Analysis - Loyal User Reward</u>

1) <u>Task:</u> Identify the five oldest users on Instagram from the provided database.

<u>SQL Command:</u>

SELECT * FROM users
SELECT username,created_at from users order by created_at limit 5;

<u>SQL Result:</u>

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

- **<u>Findings:</u>**

  The query identified the five oldest users on the platform based on their registration dates. These users have been with Instagram since its early stages.

- **<u>Insight:</u>**

  These users are the most loyal and can be targeted for loyalty rewards or exclusive offers to retain them and maintain long-term engagement.

A) Marketing Analysis - Inactive User Engagement

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

SQL Command: select * from photos,users;
                             SELECT username
                             FROM users
                             LEFT JOIN photos
                             ON users.id=photos.user_id
                             WHERE photos.id IS NULL;

SQL Result:

username
Aniya_Hackett
Kasandra_Homenick
Jadyn81
Rocio33
Maxwell.Halvorson
Tierra.Trantow
Pearl7
Ollie_Ledner37
Mckenna17
David.Osinski47
Morgan.Kassulke
Linnea59
Duane60
Julien_Schmidt
Mike.Auer39
Franco_Keebler64
Nia_Haag
Hulda.Macejkovic
Leslie67
Janelle.Nikolaus81
Darby_Herzog
Esther.Zulauf61
Bartholome.Bernhard
Jessyca_West
Esmeralda.Mraz57
Bethany20

Result 14 ✕

- **Findings:**
  The query returned a list of users who have signed up but have not posted any content. This segment represents inactive users.

- **Insight:**
  These users present a re-engagement opportunity. Marketing teams can reach out to them with special promotions or encourage them to share content, thus increasing platform activity.

A) Marketing Analysis - Contest Winner Declaration

3) **Task:** Determine the winner of the contest and provide their details to the team.

SQL Command: select * from likes, photos, users;
select likes.photo_id,users.username, count(likes.user_id) as nooflikes
from likes inner join photos on likes.photo_id=photos.id
inner join users on photos.user_id=users.id group by
likes. photo_id,users.username order by nooflikes desc;

SQL Result: Zack_Kemmer93 is winner (Highest Likes)

photo_id	username	nooflikes
145	Zack_Kemmer93	48
127	Malinda_Streich	43
182	Adelle96	43
123	Seth46	42
30	Presley_McClure	41
52	Annalise.McKenzie16	41
61	Delpha.Kihn	41
147	Meggie_Doyle	41
174	Elenor88	41
192	Kathryn80	41
256	Javonte83	41
13	Harley_Lind18	40
97	Irwin.Larson	40
153	Aurelie71	40
161	Cesar93	40
244	Damon35	40
44	Alexandro35	39
62	Kenneth64	39
66	Eveline95	39
100	Yvette.Gottlieb91	39
107	Yazmin_Mills95	39
110	Kelsi26	39
118	Janet.Armstrong	39
119	Janet.Armstrong	39
144	Zack_Kemmer93	39
180	Florence99	39
199	Donald.Fritsch	39

Result 16 ✕

- **Findings:**
  The user with the most likes on a single post was identified, along with their specific post details.

- **Insight:**
  Announcing the winner of the contest helps drive further engagement and encourages other users to compete in similar campaigns. Recognizing the most-liked posts can increase user participation.

A) Marketing Analysis - Hashtag Research

4) Task: Identify and suggest the top five most commonly used hashtags on the platform.

SQL Command:

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

SQL Result:

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

- **Findings:**
  The query identified the top five hashtags that are used most frequently by users on Instagram. These hashtags are trending and could provide high visibility for posts.

- **Insight:**
  These trending hashtags are valuable for brand partnerships. Partner brands can incorporate these hashtags to increase their content's visibility and engagement with the target audience.

A) <u>Marketing Analysis - Ad Campaign Launch</u>

5) <u>Task:</u> Determine the day of the week when most users register on Instagram. Provide insights on when to schedule an ad campaign.

<u>SQL Command:</u>

select * from users;
select DATE_FORMAT((created_at), '%W') as dayy,count(username) from users group by 1 order by 2 desc;

<u>SQL Result:</u>

dayy	count(username)
Thursday	16
Sunday	16
Friday	15
Tuesday	14
Monday	14
Wednesday	13
Saturday	12

- **Findings:**
  The query revealed the day of the week when the highest number of users registered on Instagram. This day can be considered ideal for launching new ad campaigns.

- **Insight:**
  Scheduling ads on the most active day will likely lead to a higher reach and engagement, as more users are actively registering and engaging with content.

B) **Investor Metrics - User Engagement**

1) **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.

SQL Command:

```
select * FROM PHOTOS,USERS;
with base as(
select u.id as userid, count(p.id) as photoid from users u left join photos p on p. user_id=u.id
group by u.id)
select sum(photoid) as totalphotos, count(userid) as total_users, sum(photoid)/count(userid)
as photoperuser
from base;
```

SQL Result:

Result Grid	Filter Rows:	Export:	Wrap Cell Content:
totalphotos	**total_users**	**photoperuser**	
257	100	2.5700	

- **Findings:**
  The average number of posts per user was calculated along with the total photos on Instagram. This gives an indication of how actively users are contributing content.

- **Insight:**
  A higher average number of posts per user indicates a highly engaged user base. This metric also helps investors assess the platform's overall content production and engagement levels.

B) Investor Metrics - Bots & Fake Accounts

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

SQL Command:

```
select * from users,likes;
with base as(
select u.username, count(l.photo_id) as likess from likes l inner join users u on u.id=l.user_id
group by u.username)
 select username, likess from base where likess=(select count(*) from photos) order by
username;
```

SQL Result:

username	likess
Aniya_Hackett	257
Bethany20	257
Duane60	257
Jaclyn81	257
Janelle.Nikolaus81	257
Julien_Schmidt	257
Leslie67	257
Maxwell.Halvorson	257
Mckenna17	257
Mike.Auer39	257
Nia_Haag	257
Ollie_Ledner37	257
Rocio33	257

- **Findings:**
  The query identified users who have liked every single photo on the platform, which is an unusual behaviour typically associated with bot activity.

- **Insight:**
  Identifying and removing potential bots ensures the integrity of the platform. This also improves user experience by maintaining a community of real users, making Instagram a more trusted platform.

Conclusion

This project, conducted as part of a live experience at Trainity, has provided valuable insights into the Instagram user base, engagement trends, and potential areas of improvement. The use of SQL and MySQL Workbench proved the tool for analysing key information to answer the marketing, product, and investor teams' specific queries.

Key Takeaways

a) **More Rewarding Loyalty Programs:** This was identified from the most loyal users for the foundation of exclusive rewards and retention strategies.

b) **Re-Activation Opportunities:** Identified inactive users, providing a ready-made focus for promotional activities to encourage activity.

c) **Content Performance:** Known contest winners and popular hashtags, allowing the marketing team to stimulate user engagement and facilitate brand collaborations.

d) **Optimal Launch Date for Advertising:** Identified registration trends for users to recommend the most effective days for an ad launch, thus enabling campaigns to be launched more efficiently.

e) **Investor Engagement Metrics:** Average user activity calculated and potential bot accounts identified for upholding data integrity and trust.

This project demonstrated skills in:

f) Writing SQL queries to extract actionable insights-optimizing them for efficiency.

g) Analysis of user data that can be used to make business decisions by different teams.

h) Use of data analysis in real-world scenarios for potential business impact.

i) Presentation of findings in a structured, visually appealing fashion.

END