INSTAGRAM USER ANALYTICS

Project Description :

This project aims to analyzes Instagram's user engagement metrics to detect potential bot or fake accounts, which is crucial for understanding investor insights. It involves calculating the average posts per user and identifying suspicious activities such as excessive liking. Through data-driven analysis, the project aims to offer clear insights into user behavior and platform integrity, assisting stakeholders in making informed decisions about Instagram's operational dynamics.

Approach :

Database Creation : Created and inserted the values in the database using the DDL & DML SQL queries provided by the product manager (as per project) in the MySQL database using MySQL workbench.

Extraction of Insights: After setting up the database, ran SQL queries in MySQL Workbench to extract and analyze information. These queries revealed the most followed users and the posts with the most likes .

Tech-Stack Used :

The project uses MySQL Workbench for its user-friendly interface and powerful database management tools. MySQL Server is chosen for its reliability and efficiency in handling large

datasets. SQL is utilized for robust querying capabilities to extract and analyze data. Data visualization tools are employed to create clear and insightful visual reports.

Insights :

Working on this project provided valuable insights into Instagram user behavior and platform dynamics:

A) Marketing Analysis:

1. Loyal User Reward :

Identified the five oldest users based on registration date, indicating long-term user loyalty and potential for targeted rewards.

Code :

```
/*1.Find the 5 oldest of the Instagram from the database provided */

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

The five oldest users based on registration date :

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. Inactive User Engagement :

Found users who have never posted a single photo, enabling targeted promotional strategies to encourage activity among inactive users.

Code :

```
/*2.Find the users who have never posted a single photo on Instagram*/

SELECT u.id, username FROM users AS u
LEFT JOIN photos AS ph
ON u.id = ph.user_id
WHERE ph.image_url IS NULL ;
```

The users who have never posted a single photo :

id	username
5	Aniya_Hackett
7	Kasandra_Homenick
14	Jadyn81
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

3. Contest Winner Declaration :

Determined the user with the most likes on a single photo, facilitating fair contest adjudication and prize allocation.

Code :

```
/* 3. Identify the winner of the contest and provide their details to the team*/

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

The user with the most likes on a single photo :

user_id	username	image_url	photo_id	total_likes
52	Zack_Kemmer93	https://jarret.name	145	48

4. Hashtag Research :

Identified the top five most commonly used hashtags, crucial for optimizing brand visibility and reach through effective hashtag usage.

Code :

```
/*4.Identify and suggest the top 5 most commonly used hashtag on the platform*/

SELECT t.id ,t.tag_name,COUNT(pt.photo_id) AS hashtags
FROM tags AS t
JOIN  photo_tags AS pt  ON t.id = pt.tag_id
GROUP BY t.id
ORDER BY hashtags DESC LIMIT 5;
```

The top five most commonly used hashtags :

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

5. Ad Campaign Launch :

Discovered the best day of the week for user registrations, providing strategic insights for scheduling ad campaigns to maximize user engagement.

Code :

```
/*5. What day of the week do most users register on?
Provide insights on when to schedule an ad campaign*/

SELECT DAYNAME(created_at) AS day_of_the_week, COUNT(*) AS total
FROM users GROUP BY day_of_the_week ORDER BY total DESC LIMIT 2;
```

Highest registration day :

day_of_the_week	total
Thursday	16
Sunday	16

B) Investor Metrics :

1. User Engagement Metrics :

Calculated the average posts per user and assessed overall user activity through total photos divided by total users, gauging platform health and engagement levels.

Code:

```
/*6.i)calculating the average number of posts per user */

SELECT COUNT(id) AS total_posts ,COUNT(DISTINCT user_id) AS total_users ,
COUNT(id)/COUNT(DISTINCT user_id) AS average_post_per_user FROM photos;
```

The average posts per user :

total_posts	total_users	average_post_per_user
257	74	3.4730

Code:

```
/*6.ii) Total number of photos divide by the total number of users*/

SELECT COUNT(p.id) AS total_posts, COUNT(DISTINCT u.id) AS total_users,
(COUNT(p.id)/COUNT(DISTINCT u.id)) AS average
FROM users AS u
LEFT JOIN photos AS p ON u.id = p.user_id;
```

Assessed overall user activity through total photos divided by total users :

total_posts	total_users	average
257	100	2.5700

2. Bots & Fake Accounts :

Identified potential bots through unusual activity patterns such as liking every single photo, highlighting issues that may affect user trust and platform credibility.

Code :

```
/*7. Identify the users who have liked every single post */

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

The Potential bots through unusual activity patterns such as liking every single photo :

id	username	number_of_likes
5	Aniya_Hackett	257
14	Jadyn81	257
21	Rocio33	257
24	Maxwell.Halvorson	257
36	Ollie_Ledner37	257
41	Mckenna17	257
54	Duane60	257
57	Julien_Schmidt	257
66	Mike.Auer39	257
71	Nia_Haag	257
75	Leslie67	257
76	Janelle.Nikolaus81	257
91	Bethany20	257

Results :

This project provided insights into Instagram user behavior, such as posting frequency and popular hashtags. It guided decisions on rewarding loyal users and timing ad campaigns based on user sign-up patterns. Identification of potential fake accounts highlighted the importance of platform security. Overall, the project underscored the significance of data-driven decisions in maintaining a trustworthy and effective Instagram platform.