Instagram User Analytics

Project by- Ritika chaudhary

Project Description

The project aims to analyze user behaviour and provide insights for the marketing and investor teams of Instagram. By leveraging the provide database, the analysis will address various aspects such as rewarding loyal users, encouraging inactive users, declaring contest winners, researching popular hashtags, determining he best day to launched campaigns, assessing user engagement, and identifying potential bot accounts.

Approach

The analysis will be conducted using SQL queries on the Instagram database. The database contains tables for users, photos, comments, likes, follows, tags and more. By combining these tables and performing appropriate joins, aggregations and calculations, we can derive the necessary insights to answer the questions posed by the marketing and investor teams.

Tech-Stack Used

For this project, the SQL language will be used to interact with the provided Instagram database. Various SQL queries will be written to extract the required information from the database. The choice of the SQL database management system (DBMS) is left open-ended, as SQL is a standard language that can be executed on different platforms such as MySQL, PostgreSQL, or SQLite.
This approach was executed by using MySQL software.

INSIGHTS

		Rewarding Most Loyal Users
		Remind Inactive Users To Start Posting
01	**Marketing**	Declaring Contest Winner
		Hashtag Researching
		Launch AD Campaign
02	**Investor**	User Engagement
	Metrics	Bots & Fake Accounts

MARKETING

Rewarding Most Loyal Users

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

Task : Find the 5 oldest users of the Instagram from database provided

By identifying the 5 oldest users on Instagram,

the marketing team can reward them for their

loyalty, potentially through special promotions

or exclusive features.

```sql
select *from users
order by created_at
limit 5;
```

Remind Inactive Users to Start Posting

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

Identifying users who have never posted a photo on Instagram allows the marketing team to target them with promotional emails or campaigns to encourage their engagement.
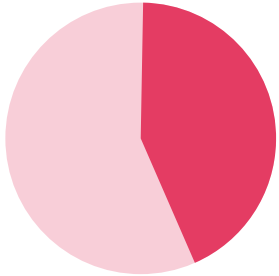
1 billion
Active users

```sql
SELECT u.id, u.username
FROM users u
LEFT JOIN photos p ON u.id = p.user_id
WHERE p.id IS NULL;
```

Declaring Contest Winner

By determining the user who received the most likes on a single photo, the contest winner can be declared, leading to increased engagement and excitement among users.

```sql
SELECT u.id, u.username, p.image_url, COUNT(l.user_id) AS like_count
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, p.id
ORDER BY like_count DESC
LIMIT 1;
```

	id	username	image_url	like_count
▶	52	Zack_Kemmer93	https://jarret.name	48

%

Likes

Hashtag Researching

Approach

```
# 5 most commonly used hashtags
SELECT t.tag_name, COUNT(*) AS count
FROM tags t
JOIN photo_tags pt ON t.id = pt.tag_id
GROUP BY t.tag_name
ORDER BY count DESC
LIMIT 5;
```

Result

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

Identifying the top 5 most commonly used hashtags on Instagram can help a partner brand optimize their post reach and engagement by using popular hashtags. The followings hashtags are the most commonly used hashtags on the platform

Launch AD Campaign

By analyzing the registration patterns of the users, the team can determine the best day to schedule an AD campaign for the maximum impact and reach.

Best days to post on IG

Mon	Tue	Wed	Thu	Fri	Sat	Sun

Best day

```
# launch AD campaign
SELECT DAYNAME(created_at) AS registration_day, COUNT(*) AS total_registrations
FROM users
GROUP BY registration_day
ORDER BY total_registrations DESC
LIMIT 1;
```

	registration_day	total_registrations
▶	Thursday	16

User Engagement

Calculating the average number of posts per user provides insights into user activity levels on instagram. Analyzing the total number of photos in relation to the toatal number of users helps assess user participation

01

```sql
SELECT COUNT(*) / COUNT(DISTINCT user_id) AS average_posts_per_user
FROM photos;
```

average_posts_per_user
3.4730

02

```sql
#total number of photos on Instagram/total number of users
SELECT COUNT(*) / (SELECT COUNT(*) FROM users) AS photos_per_user_ratio
FROM photos;
```

photos_per_user_ratio
2.5700

Bots & Fake Accounts

The investors want to know if the platform is crowded with fake and dummy accounts.
Identifying users who have liked every single photo on the site helps detect potential bot or fake accounts exhibit abnormal behaviour.
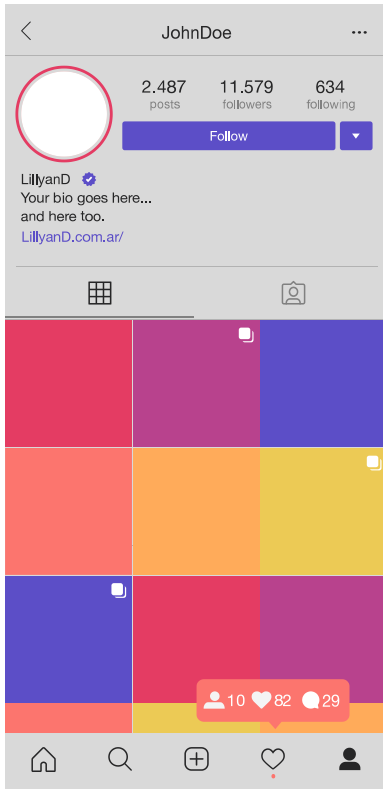

```sql
SELECT u.id, u.username
FROM users u
INNER JOIN likes l ON u.id = l.user_id
GROUP BY u.id
HAVING COUNT(DISTINCT l.photo_id) = (SELECT COUNT(*) FROM photos);
```

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

RESULTS

- This project has provided me with hands-on experience in data analysis, query optimization and insights generation.
- It has helped me enhance my technical skills and business acumen.
- While working on this project, I have achieved several things that have helped me in learning and understanding of data analysis and providing actionable insights.