Instagram
User Engagement Analysis Using SQL

CREATED BY : SAKSHI DIVE

Project Overview

This project focuses on analyzing user interactions and engagement on Instagram to extract actionable insights that can drive business growth. User analysis tracks how individuals interact with the app helping businesses make data-driven decisions across various teams such as marketing , product and development to enhance user experience and drive business growth.

Using **SQL** and **MySQL Workbench**, this analysis answers key management questions providing valuable insights to shape the app's future direction.

Objectives

▶ The primary goal of this project is to utilize SQL to analyze Instagram user data and derive meaningful insights that can support strategic decision-making. Specifically, this project aims to answer following questions:

1. Identify the **5 Most Active Users** on the platform based on their posting activity.

2. Determine which **User** received the **Highest Number of Likes** on a single photo.

3. Analyze the **Tags** associated with the **Most Liked Photo** to identify trends in user preferences.

4. Discover and recommend the **5 most commonly used Tags** on the platform to enhance engagement.

5. Examine user registration trends to determine the **Day of the Week with the Highest Registrations.**

6. Identify users who have **Never posted** to understand **Inactive or Passive user** behavior.

7. Detect users exhibiting **Bot-like or Fake Account Behavior** through patterns of excessive or automated engagement, such as liking every photo.

Dataset and Tools Used

▶ **Dataset** : The dataset "**instagram_db**" consist of 6 tables

1. **Users :** Information about users
2. **Photos:** Data about photos posted by users
3. **Likes:** Records of which users liked which photos
4. **Follows:** Follower-followee relationships
5. **Tags:** Information about tags
6. **Photo Tags**: Tags associated with photos

▶ **Tools Used** :

SQL : Querying and Analyzing the data

 MySQL Workbench : Managing database and executing queries.

5 Most Active Users on the platform based on posting photos.

Query

```sql
select u.id , username, count(p.image_url) as No_of_Photos
from photos p
left join users u on p.user_id = u.id
group by u.id , u.username
order by count(p.image_url) desc limit 5 ;
```

Output

id	username	No_of_Photos
23	Eveline95	12
88	Clint27	11
59	Cesar93	10
86	Delfina_VonRueden68	9
58	Aurelie71	8

User who received Highest Number of Likes on a Single Photo

Query

```sql
select u.username , l.photo_id from likes l
left join photos p on l.photo_id = p.id
left join users u on p.user_id = u.id
group by u.username , l.photo_id
order by count(l.user_id) desc limit 1;
```

Output

username	photo_id
Zack_Kemmer93	145

Tags associated with Most Liked Photo

Query

```sql
select t.tag_name, pt.photo_id from tags t
join photo_tags pt on t.id = pt.tag_id
where pt.photo_id = 145;
```

Output

	tag_name	photo_id
▸	fun	145
	party	145
	concert	145
	drunk	145
	smile	145

5 Most commonly used Tags

Query

```
select t.tag_name from tags t
left join photo_tags pt on t.id = pt.tag_id
group by t.tag_name
order by count(pt.tag_id) desc limit 5;
```

Output

tag_name
smile
beach
party
fun
concert

Day of the Week with Highest Registrations

Query

```sql
select dayname(created_at) as dayy, count(username) as No_of_Registrations
from users
group by dayy
order by No_of_Registrations desc limit 2 ;
```

Output

dayy	No_of_Registrations
Thursday	16
Sunday	16

Users who have Never Posted to understand **Inactive or Passive User** behavior.

Query

Output

```
select u.username from users u
left join photos p on u.id = p.user_id
where p.image_url is null
order by username ;
```

username
Aniya_Hackett
Bartholome.Bernhard
Bethany20
Darby_Herzog
David.Osinski47
Duane60
Esmeralda.Mraz57
Esther.Zulauf61
Franco_Keebler64
Hulda.Macejkovic
Jaclyn81
Janelle.Nikolaus81
Jessyca_West

username
Julien_Schmidt
Kasandra_Homenick
Leslie67
Linnea59
Maxwell.Halvorson
Mckenna17
Mike.Auer39
Morgan.Kassulke
Nia_Haag
Ollie_Ledner37
Pearl7
Rocio33
Tierra.Trantow

Users exhibiting Bot-like or Fake Account behavior

Query

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

Output

username	id
Aniya_Hackett	5
Jaclyn81	14
Rocio33	21
Maxwell.Halvorson	24
Ollie_Ledner37	36
Mckenna17	41
Duane60	54
Julien_Schmidt	57
Mike.Auer39	66
Nia_Haag	71
Leslie67	75
Janelle.Nikolaus81	76
Bethany20	91

Insights

▶ **User Engagement**: Identified the **5 Most Active Users** contributing significantly to platform engagement, suggesting potential influencers or brand ambassadors.

▶ **Content Trends**: **Smile, Fun, Party, Concert** and **Drunk** are the tags associated with most liked photo whereas **Smile, Beach, Party, Fun** & **Concert** are 5 most commonly used tags which highlights trends in user preferences which can be actionable insights for content optimization.

▶ **User Behavior**: Detected **13 Users** exhibiting **Bot-like or Fake Account Behavior** enabling the platform to take action against fake accounts which can help to maintain the platform's integrity.

▶ **Registration Trends**: **Thursday** and **Sunday** are the peak days for registrations informing targeted marketing campaigns and opportunities for growth.

▶ **Inactive Users**: Identified **26 Inactive Users**, revealing opportunities for reactivation to encourage user sig-ups.

Recommendations

- **Leverage Influencers**: Engage the most active users through exclusive features, rewards, or collaborations to promote content and maintain their activity and leverage their influence..

- **Optimize Content Strategy**: Encourage the use of popular tags and themes to boost engagement.

- **Mitigate Fake Accounts**: Implement stricter bot detection mechanisms or verification processes to maintain platform authenticity.

- **Enhance User Activation**: Design campaigns targeting inactive users to encourage participation reach out to them with engagement prompts, tutorials, or incentives to encourage their activity.

- Schedule promotional ad campaigns and onboarding processes around the peak registration days.

Impacts

▶ **Actionable Business Insights**: Insights derived empowers **Marketing, Product, and Development teams** to launch targeted campaigns, introduce new features, and enhance user experience, driving business growth.

▶ **Improved User Engagement**: Detailed analysis of user activity and tag trends help to **optimize content strategies**, boosting user interaction and engagement.

▶ **Enhanced Platform Trust**: Detecting and addressing **bot-like behavior** ensures accurate engagement metrics, fostering **trust among users.**

▶ **Higher User Satisfaction**: Data-driven content optimization makes posts more relevant, increasing **engagement rates** and user satisfaction.

▶ **Strategic Growth Opportunities**: Leveraging registration trends and influencer partnership supports **marketing campaigns**, driving **user acquisition** and retention.

Conclusion

This project demonstrates how SQL can extract actionable insights to optimize engagement, content strategies, and platform integrity. By analyzing user behavior and trends, we identified growth opportunities and highlighted the critical role of data-driven decisions in enhancing user experience and supporting business success.

This analysis showcases my ability to use SQL to deliver meaningful insights that supports data-driven decisions and business growth.

Thank You !