

# Instagram User Analytics Project

## Project Description

This project focuses on analyzing user interactions and engagement on Instagram using SQL. The aim is to provide insights that can inform marketing strategies, product development, and investor evaluations. Various SQL queries are used to extract data related to user activity, popular content, and user engagement metrics.

## Approach

The project approach is structured into two main parts: Marketing Analysis and Investor Metrics. Each part includes specific tasks such as identifying loyal users, calculating average engagement levels, and detecting potential bot accounts. MySQL Workbench is used to create and manage the database, execute SQL queries, and analyze the results. Each SQL query is written to maximize efficiency and accuracy, ensuring the insights are both comprehensive and actionable.

## Tech-Stack Used

MySQL Workbench was the primary tool for this project. Its intuitive interface and powerful query capabilities made it ideal for analyzing large datasets and generating complex SQL queries. Version used: MySQL Workbench 8.0.

## SQL Tasks and Queries

### A) Marketing Analysis

#### *Identify the Five Oldest Users on Instagram*

The query below identifies the five users who registered earliest on Instagram, providing insights into loyal users who have been with the platform the longest.

```
SELECT username, created_at  
FROM users  
ORDER BY created_at ASC
```

LIMIT 5;

Output :

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

### *Identify Inactive Users*

This query identifies users who have never posted a photo, helping the marketing team target users who may benefit from re-engagement.

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

Output :

Aniya_Hackett
Bartholome.Bernhard
Bethany20
Darby_Herzog
David.Osinski47
Duane60
Esmeralda.Mraz57
Esther.Zulauf61
Franco_Keebler64

Hulda.Macejkovic  
Jaclyn81  
Janelle.Nikolaus81  
Jessyca\_West  
Julien\_Schmidt  
Kasandra\_Homenick  
Leslie67  
Linnea59  
Maxwell.Halvorson  
Mckenna17  
Mike.Auer39  
Morgan.Kassulke  
Nia\_Haag  
Ollie\_Ledner37  
Pearl7  
Rocio33  
Tierra.Trantow

### *Identify Contest Winner*

The contest winner is determined based on the photo with the most likes, which identifies the user and their photo that received the highest engagement.

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

Output :

52 Zack\_Kemmer93 145 <https://jarret.name> 48

### *Identify Top Five Hashtags*

This query provides the top five hashtags, helping partner brands understand the most popular tags for maximizing reach.

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

Output :

smile 59

beach 42

party 39

fun 38

concert 24

### *Determine Optimal Day for Ad Campaign*

This query finds the day with the highest user registrations, which can help in scheduling ad campaigns for maximum impact.

```
SELECT DAYNAME(created_at) AS registration_day, COUNT(id) AS count_registered
FROM users
GROUP BY registration_day
ORDER BY count_registered DESC
LIMIT 1;
```

Output :

Thursday 16

## B) Investor Metrics

### *Calculate Average Posts per User*

This query calculates the average number of posts per user, giving investors insights into the general engagement level on the platform.

```
SELECT AVG(photo_count) AS avg_posts_per_user
FROM (
    SELECT user_id, COUNT(id) AS photo_count
    FROM photos
    GROUP BY user_id
) AS user_photos;
```

Output :

3.4730

### *Calculate Photos per User Ratio*

This query calculates the ratio of total photos to total users on Instagram, providing a high-level view of engagement.

```
SELECT
    (SELECT COUNT(*) FROM photos) / (SELECT COUNT(*) FROM users) AS
    photos_per_user_ratio;
```

Output :

2.5700

### *Identify Potential Bots*

The query identifies potential bot accounts by finding users who have liked every photo, which may indicate automated activity.

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

Output :

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

## Insights

The analysis provided insights on user loyalty, engagement, and common interests (hashtags). Notably, identifying inactive users and popular hashtags can inform re-engagement campaigns and content strategy. Also, the analysis of average posts per user and bot detection helps investors understand user authenticity and engagement levels. These insights can influence future development and marketing strategies on Instagram.

## Results

Through this project, I gained valuable experience in SQL querying and database analysis. The project provided actionable insights on user engagement, popular content types, and potential growth areas for Instagram. The SQL analysis demonstrated my ability to handle real-world data and present meaningful findings for strategic decision-making.