# Instagram User Analytics

## Project Description:

This study delves into the valuable user interaction data from the Instagram app. Through data analysis, we aim to extract insights to guide future product development and marketing strategies. By examining user behavior patterns, we can spot trends, understand preferences across demographics, and enhance the overall user experience to retain engagement. This project offers insights into popular content types and uncovers how users engage with the app's features. Understanding user behavior is crucial for shaping Instagram's future and ensuring success in the dynamic social media landscape.

## Project Approach:

**1. Explore user data:** Understand demographics, post interactions, app time, and content preferences using MySQL Workbench.

**2. Formulate queries:** Collaborate with the product manager to translate key questions into targeted SQL queries that address user engagement, demographics, content types, regional variations, and feature usage.

**3. Analyze data:** Uncover trends in user activity to identify ideal posting times, content formats for specific user groups, and feature optimization opportunities.

**4. Share insights:** Present actionable findings through reports and dashboards to inform marketing campaigns, feature development, and user experience improvements.

# Tech Stack Used:

 **MySQL Workbench:** A visual database management tool for querying and analyzing Instagram user data stored in a relational database.

 **SQL (Structured Query Language):** A programming language used to interact with relational databases, allowing for data extraction, manipulation, and analysis.

# Project Insights:

The analysis has the potential to reveal valuable insights into user behavior on Instagram, potentially uncovering:

a) Marketing Analysis
b) Investor Metrics

# A) Marketing Analysis

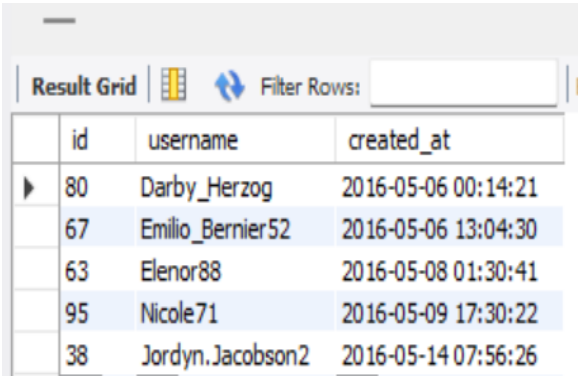
## 1) Loyal User Reward:

Identifying the 5 oldest Instagram users from the given database:

### *Query:*

```
SELECT *
FROM users
ORDER BY created_at
LIMIT 5;
```

### *Output:*




| 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 |

### *Result:*

Hence the 5 oldest Instagram users are - a) Darby_Herzog, b) Emilio_Bernier52, c) Elenor88, d) Nicole71 and e) Jordyn.Jacobson2

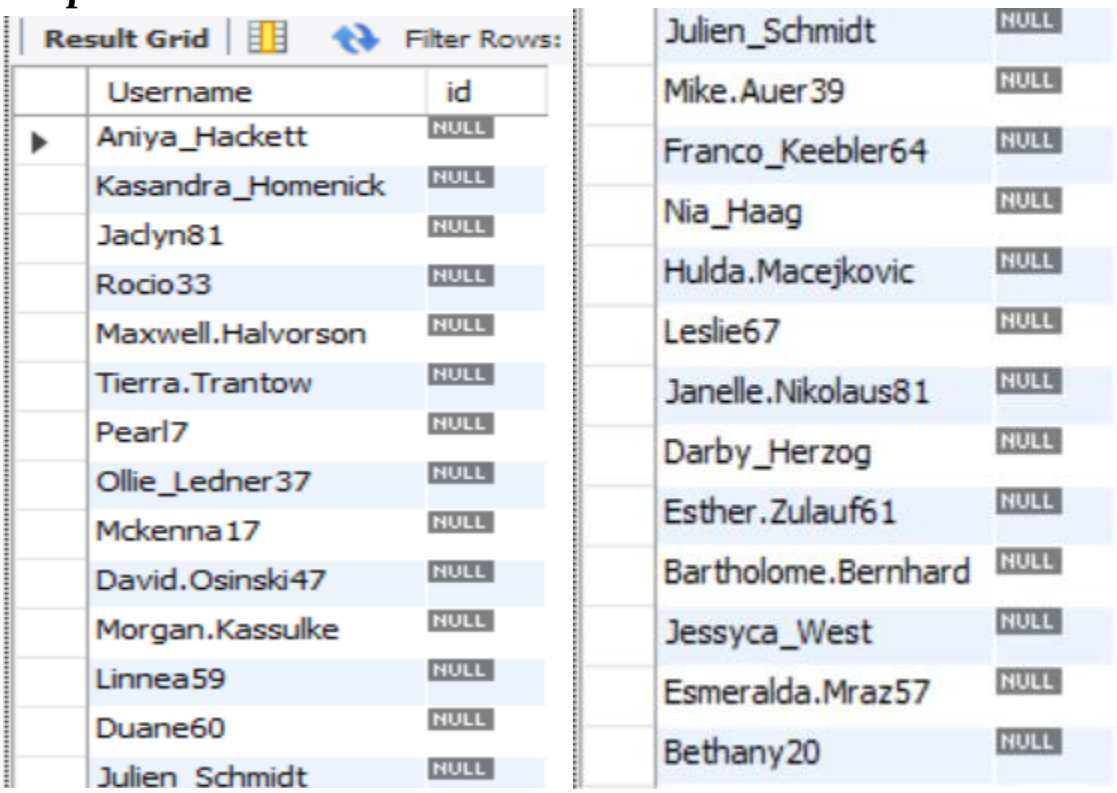
## 2) Inactive User Engagement:

Identifying users who have never posted a single picture on Instagram

### *Query:*

```
SELECT users.username
AS Username, Photos.id
FROM users
LEFT JOIN photos
ON users.id = photos.user_id
WHERE photos.id IS NULL;
```

### *Output:*



Result Grid | Filter Rows:

| Username | id |
|---|---|
| Aniya_Hackett | NULL |
| Kasandra_Homenick | NULL |
| Jaclyn81 | NULL |
| Rocio33 | NULL |
| Maxwell.Halvorson | NULL |
| Tierra.Trantow | NULL |
| Pearl7 | NULL |
| Ollie_Ledner37 | NULL |
| Mckenna17 | NULL |
| David.Osinski47 | NULL |
| Morgan.Kassulke | NULL |
| Linnea59 | NULL |
| Duane60 | NULL |
| Julien_Schmidt | NULL |
| Julien_Schmidt | NULL |
| Mike.Auer39 | NULL |
| Franco_Keebler64 | NULL |
| Nia_Haag | NULL |
| Hulda.Macejkovic | NULL |
| Leslie67 | NULL |
| Janelle.Nikolaus81 | NULL |
| Darby_Herzog | NULL |
| Esther.Zulauf61 | NULL |
| Bartholome.Bernhard | NULL |
| Jessyca_West | NULL |
| Esmeralda.Mraz57 | NULL |
| Bethany20 | NULL |

### *Result:*

The above users with id = NULL are the users who have never posted any picture on the platform.

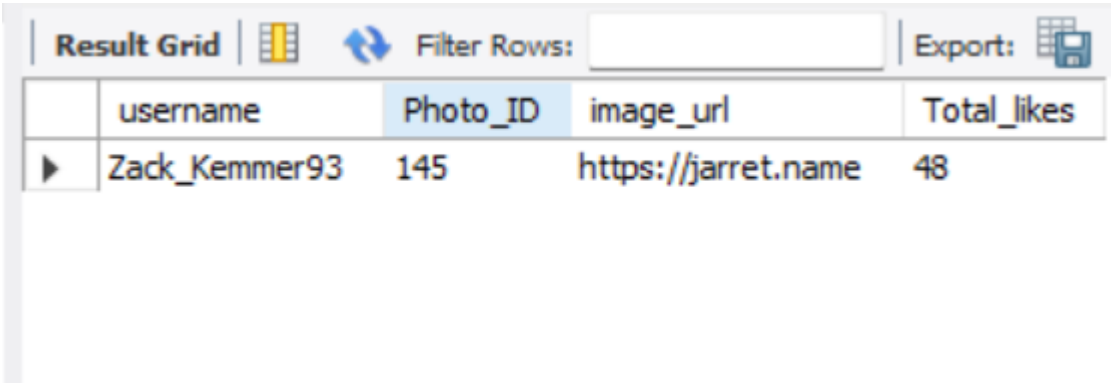
## 3) Contest Winner Declaration:

Determining the winner of the contest and providing their details to Team

### *Query:*

```
SELECT
username,
photos.id AS Photo_ID,
 photos.image_url,
COUNT(*) AS Total_likes
FROM photos
INNER JOIN likes
  ON likes.photo_id = photos.id
INNER JOIN users
  ON photos.user_id = users.id
GROUP BY photos.id
ORDER BY Total_likes DESC
LIMIT 1;
```

### *Output:*




### *Result:*

Zack_Kemmer93 is the winner of the Contest with Total_likes of 48

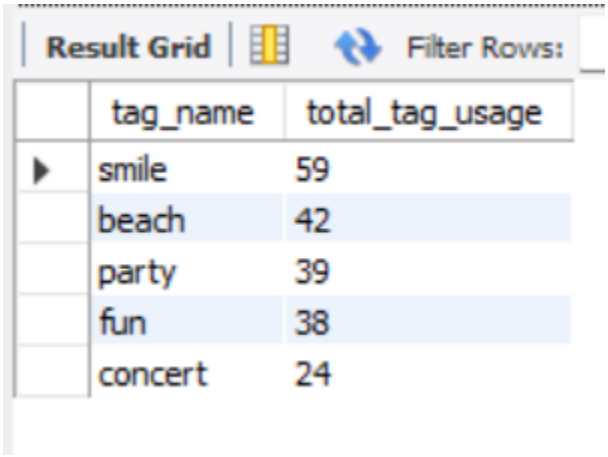
## 4) Hashtag Research:

Identifying and suggesting top 5 hashtags used on the platform

### *Query:*

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

### *Output:*




| tag_name | total_tag_usage |
|---|---|
| smile | 59 |
| beach | 42 |
| party | 39 |
| fun | 38 |
| concert | 24 |

### *Result:*

Top 5 popular hashtags on Instagram are – smile, beach, party, fun, and concert.

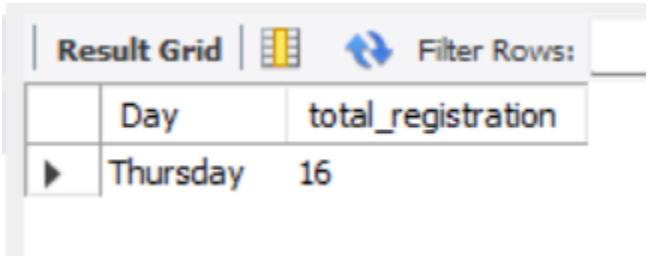
# 5) AD Campaign Launch:

Determining the day of the week when the most users register on instagram. Provide an insight so as to when to launch an ad campaign.

## *Query:*

```
SELECT
dayname(created_at)
AS Day ,
COUNT(*) AS total_registration
FROM users
GROUP BY Day
ORDER BY total_registration DESC
LIMIT 1;
```

## *Output:*




## *Result:*

Best day according to the database provided to launch an ad campaign is – Thursday with 16 registration on that day.

# B) Investor Metrics

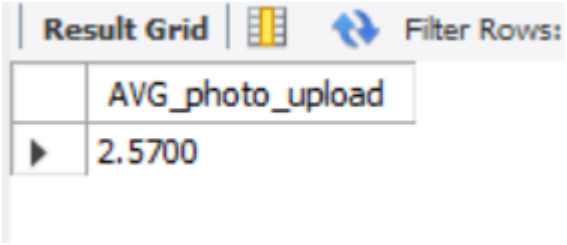
## 1) User Engagement:

Calculating the average number of post per user on Instagram. Also finding the total number of posts on Instagram divided by total number of users

***Query:***

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

***Output:***




***Result:***

The average number of post uploaded per user is 2.5700

## 2) Bots and Fake Accounts

Identifying users (Potential bots) who have liked every picture on Instagram as it is typically not possibe by any user.

### *Query:*

```
SELECT users.id,username, COUNT(users.id)
AS total_likes_given_by_user

FROM users
INNER JOIN likes ON users.id = likes.user_id
GROUP BY users.id
HAVING total_likes_given_by_user = (SELECT COUNT(*)
FROM photos);
```

### *Output:*

Result Grid | Filter Rows: | Export:

| id | username | total_likes_given_by_user |
|---|---|---|
| 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 |

### *Result:*

The users with id 5, 14, 21, 24, 36, 41, 54, 57, 66, 71, 75, 76, and 91 may be potential bots.

# Conclusion

The exploration of Instagram user data through SQL queries and MySQL Workbench was a successful project. Valuable insights were gained from this analysis, providing guidance to various teams at Instagram, such as marketing, product development, and investor relations. This project demonstrates the importance of making decisions based on data to influence the future direction of one of the worlds leading social media platforms.

Through this project, I gained a deeper comprehension of user behavior and interaction on Instagram. By utilizing SQL effectively for data analysis, I played a role in the strategic decision-making process of the platform.