

# INSTAGRAM USER ANALYTICS

## ❖ Project Description

This project focuses on analysing Instagram user data to provide actionable insights to various teams within the company. As a data analyst, I will be using SQL and MySQL Workbench to perform several analyses. These analyses will help the marketing team reward loyal users, engage inactive users, declare contest winners, identify popular hashtags, and determine the best day to launch ad campaigns. Additionally, I will provide metrics for investors regarding user engagement and potential fake accounts and extract useful insights from raw data, using tools, and even visualize them to increase platform's efficiency.

## ❖ Approach

**Data Preparation:** Import the Instagram dataset into MySQL Workbench and ensure the data is clean and properly structured for analysis.

### ***Marketing Analysis:***

- Loyal User Reward: Query the database to find the five oldest users based on their registration date.
- Inactive User Engagement: Identify users who have never posted a single photo by checking the post count for each user.
- Contest Winner Declaration: Find the user with the most likes on a single photo by joining user and photo tables and sorting by likes.
- Hashtag Research: Identify the top five most used hashtags by aggregating hashtag usage data.
- Ad Campaign Launch: Determine the day of the week with the highest user registrations to suggest the optimal day for launching ads.

### ***Investor Metrics:***

- User Engagement: Calculate the average number of posts per user and the total number of photos divided by the total number of users.
- Bots & Fake Accounts: Identify users who have liked every single photo on the site by performing a comprehensive join and count operation on likes.

## ❖ Tech-Stack Used

- MySQL Workbench v8.0.30.0 : Chosen for its powerful SQL querying capabilities and ease of use for database management and analysis, its simple setup and GUI.

## ❖ Insights

### **A. Marketing Analysis:**

1. Loyal User Reward: Identified the five oldest users, enabling the marketing team to reward loyalty effectively.

### **CODE:**

#### **#1 LOYAL USERS**

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

**OUTPUT:**

	id	username	created_at
	180	Darby_Herzog	2016-05-06 00:14:21
	167	180 p_Bernier52	2016-05-06 13:04:30
	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
▶*	NULL	NULL	NULL

2. Inactive User Engagement: Found users with zero posts, providing a target group for re-engagement campaigns.

**CODE:**

#2 INACTIVE USERS ENGAGEMENT

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

**OUTPUT:**

	username
▶	Aniya_Hackett
	Kassandra_Homenick
	Jadlyn81
	Rocio33
	Maxwell.Halvorson
	Tierra.Trantow
	Pearl7
	Ollie_Ledner37
	Mckenna17
	David.Osinski47
	Morgan.Kassulke
	Linnea59
	Duane60
	Julien_Schmidt
	Mike.Auer39
	Franco_Keebler64
	Nia Haas

3. Contest Winner Declaration: Determined the user with the most likes on a single photo, ensuring a transparent contest process.

**CODE:**

#3 CONTEST WINNER

```
SELECT
username,photos.id,photos.image_url,count(likes.user_id) AS total
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 DESC
LIMIT 1;
```

**OUTPUT:**

	username	id	image_url	total
▶	Zack_Kemmer93	145	https://jarret.name	48

4. Hashtag Research: Revealed the top five most popular hashtags, aiding partner brands in optimizing their reach.

**CODE:**

#4 MOST HASHTAG

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

**OUTPUT:**

	tag_name	total
▶	smile	59
	beach	42
	party	39
	fun	38
	concert	38

5. Ad Campaign Launch: Discovered the day with the highest user registrations, suggesting the best day to launch ad campaigns.

**CODE:**

#5 BEST DAY TO LAUNCH

```
SELECT
DAYNAME(created_at) AS day, count(*) as total
```

```
FROM users
GROUP BY day
ORDER by total DESC
LIMIT 1;
```

**OUTPUT:**

	day	total
▶	Thursday	32

**B. Investor Metrics:**

1. User Engagement: Calculated the average posts per user, showing engagement levels across the platform.

**CODE:**

#6 USER ENGAGEMENT

```
SELECT
```

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

**OUTPUT:**

	avg
▶	1.2850

2. Bots & Fake Accounts: Identified potential bot accounts, addressing investor concerns about platform integrity.

**CODE:**

#7 FAKE ACCOUNTS

```
SELECT user_id, COUNT(*) as num_likes
```

```
FROM likes
```

```
GROUP BY user_id
```

```
HAVING num_likes=(SELECT COUNT(*) FROM photos);
```

```
SELECT u.username, COUNT(*) as num_likes
```

```
FROM users u
```

```
JOIN likes l ON u.id=l.user_id
```

```
GROUP BY u.id
```

```
HAVING num_likes=(SELECT COUNT(*) FROM photos);
```

**OUTPUT:**

	username	num_likes
▶	Aniya_Hackett	257
	Jadyn81	257
	Rocio33	257
	Maxwell.Halvorson	257
	Ollie_Ledner37	257
	Mckenna17	257
	Duane60	257
	Julien_Schmidt	257
	Mike.Auer39	257
	Nia_Haag	257
	Leslie67	257
	Janelle.Nikolaus81	257
	Bethany20	257

## Results:

The project successfully provided valuable insights and actionable recommendations for the marketing and product teams. These insights will help improve user engagement, optimize marketing strategies, and ensure transparency and integrity within the platform. The detailed analysis also addresses investor concerns by highlighting user engagement levels and identifying potential fake accounts.

## Achievements

### *Personal Achievements*

- Enhanced SQL Proficiency
- Data Analysis Expertise
- Problem-Solving Skills

### *Professional Growth*

- Technical Development: The project provided hands-on experience with SQL and MySQL Workbench, enhancing my technical capabilities and confidence in handling complex data analysis tasks.
- Analytical Thinking: It honed my analytical thinking, allowing me to approach data-driven problems with a structured and methodical mindset.
- Communication Skills: Improved my ability to communicate complex data insights in a clear and actionable manner to non-technical stakeholders.

### *Impact of the Analysis*

- Marketing Strategy: The insights I provided helped the marketing team to better engage users, target campaigns more effectively, and recognize loyal users, which can improve user retention and satisfaction.
- Product Development: The product team benefited from understanding user behaviour patterns, guiding feature enhancements, and prioritizing development efforts based on user needs.
- Investor Relations: The metrics on user engagement and the identification of potential fake accounts helped build investor confidence in the platform's integrity and growth potential.

- Strategic Decision-Making: My analysis on optimal ad launch days and popular hashtags supported strategic planning and optimized marketing efforts, leading to better resource allocation and campaign effectiveness.

### *Thoughts on the Impact of the Analysis*

The analysis had a substantial impact on both my personal and professional development. Personally, it enhanced my technical and analytical skills, preparing me for more complex data analysis projects in the future. Professionally, it provided valuable insights that could influence strategic decisions, improve user engagement, and enhance the overall user experience on Instagram.