

# Instagram User Analytics - SQL Fundamentals

## Project Description:

This project involves analyzing user interactions and engagement on the Instagram platform using SQL queries. As a data analyst, the objective is to provide valuable insights into user behavior and metrics that can help the marketing, product, and investor teams at Instagram make informed decisions. By leveraging SQL, we aim to address various business questions, from marketing analysis to investor metrics.

## Approach:

The project is divided into two main sections: Marketing Analysis and Investor Metrics. Each section involves SQL queries designed to extract data from the Instagram user database to answer specific business questions. The queries are executed in MySQL Workbench.

- **Step 1:** Understanding the data and database schema.
- **Step 2:** Writing SQL queries to address each task.
- **Step 3:** Interpreting the query outputs and summarizing insights.
- **Step 4:** Preparing the report with SQL queries and results.
- **Step 5:** Generating actionable insights from the analysis.

## Tech-Stack Used:

- **MySQL Workbench:** Used to run SQL queries and perform database operations.
- **SQL (Structured Query Language):** Used to query and analyze data efficiently from the Instagram database.
- **Google Drive:** For organizing and submitting the final report.

## SQL Tasks & Insights:

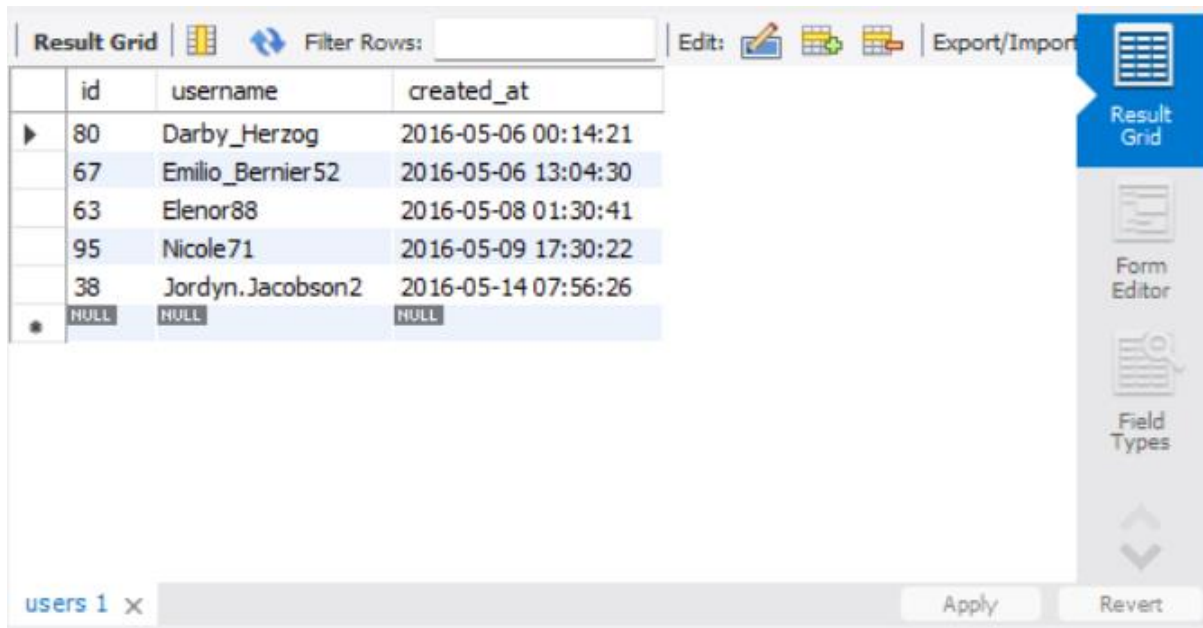
### A) Marketing Analysis:

#### 1. Loyal User Reward:

- **Query:**

```
#1 LOYAL USERS
SELECT * FROM users
ORDER BY created_at
LIMIT 5;
```

**Insight:** Identified the five oldest users who have been on the platform the longest. These users are prime candidates for loyalty rewards.



The screenshot shows a database management tool interface. At the top, there's a 'Result Grid' tab, a 'Filter Rows' input field, and buttons for 'Edit', 'Export/Import', 'Form Editor', and 'Field Types'. The main area displays a table with the following data:

|   | 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 |
| • | NULL | NULL             | NULL                |

At the bottom, there's a tab labeled 'users 1' with an 'x' icon, and buttons for 'Apply' and 'Revert'.

## 2. Inactive User Engagement:

- **Query:**

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

- **Insight:** Extracted a list of users who have never posted a photo. These users could be targeted with promotional emails to encourage posting.

|    | A                   | B | C | D | E | F | G |
|----|---------------------|---|---|---|---|---|---|
| 1  | username            |   |   |   |   |   |   |
| 2  | Aniya_Hackett       |   |   |   |   |   |   |
| 3  | Kassandra_Homenick  |   |   |   |   |   |   |
| 4  | Jaclyn81            |   |   |   |   |   |   |
| 5  | Rocio33             |   |   |   |   |   |   |
| 6  | Maxwell.Halvorson   |   |   |   |   |   |   |
| 7  | Tierra.Trantow      |   |   |   |   |   |   |
| 8  | Pearl7              |   |   |   |   |   |   |
| 9  | Ollie_Ledner37      |   |   |   |   |   |   |
| 10 | Mckenna17           |   |   |   |   |   |   |
| 11 | David.Osinski47     |   |   |   |   |   |   |
| 12 | Morgan.Kassulke     |   |   |   |   |   |   |
| 13 | Linnea59            |   |   |   |   |   |   |
| 14 | Duane60             |   |   |   |   |   |   |
| 15 | Julien_Schmidt      |   |   |   |   |   |   |
| 16 | Mike.Auer39         |   |   |   |   |   |   |
| 17 | Franco_Keebler64    |   |   |   |   |   |   |
| 18 | Nia_Haag            |   |   |   |   |   |   |
| 19 | Hulda.Macejkovic    |   |   |   |   |   |   |
| 20 | Leslie67            |   |   |   |   |   |   |
| 21 | Janelle.Nikolaus81  |   |   |   |   |   |   |
| 22 | Darby_Herzog        |   |   |   |   |   |   |
| 23 | Esther.Zulauf61     |   |   |   |   |   |   |
| 24 | Bartholome.Bernhard |   |   |   |   |   |   |
| 25 | Jessyca_West        |   |   |   |   |   |   |
| 26 | Esmeralda.Mraz57    |   |   |   |   |   |   |
| 27 | Bethany20           |   |   |   |   |   |   |

### 3. Contest Winner Declaration:

- **Query:**

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

- **Insight:** The user with the highest likes on a single photo was identified and declared the winner of the contest.

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

#### 4. Hashtag Research:

- Query:

```
#4 MOST HASHTAGS
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;
```

- **Insight:** Suggested the top five most commonly used hashtags to partner brands for maximum reach.

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

#### 5. Ad Campaign Launch:

- Query:

```
# MOST INSTAGRAM USER REGISTERED
SELECT DAYNAME(created_at) AS day, count(*) as total
FROM users
GROUP BY day
ORDER BY total DESC
LIMIT 1;
```

- **Insight:** Found that most users register on Instagram on a particular day of the week, indicating the best time to launch an ad campaign.

|   | day      | total |
|---|----------|-------|
| ▶ | Thursday | 16    |

## B) Investor Metrics:

### 1. User Engagement:

- **Query:**

```
#1 USER ENGAGEMENT
SELECT
(SELECT COUNT(*) FROM photos) / (SELECT COUNT(*) from users) AS avg;
```

- **Insight:** Calculated the average number of posts per user, providing insights into the platform's overall engagement levels.

|   | avg    |
|---|--------|
| ▶ | 2.5700 |

### 2. Bots & Fake Accounts:

- Query:**

```
#1 USER ENGAGEMENT
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);
```

- **Insight:** Identified users who liked every single photo on the platform, likely indicating bots or fake accounts.

|   | 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       |

#### Insights:

- Loyal users represent an important segment for Instagram, and rewarding them can help with retention.
- A significant number of users are inactive in terms of posting photos, which could be a target group for engagement campaigns.
- User engagement metrics indicate healthy platform activity, although potential bot accounts need further investigation.
- The use of certain hashtags can enhance brand visibility, while understanding user registration patterns can help optimize marketing efforts.

#### Result:

The analysis provided several actionable insights for Instagram's marketing and product teams. The queries helped identify key user segments such as loyal users, inactive users, and potential bots. Furthermore, insights into popular hashtags and optimal ad campaign launch times will guide future business strategies. Overall, this project demonstrated the power of SQL in analyzing large datasets and extracting valuable business insights.