Instagram User Analytics

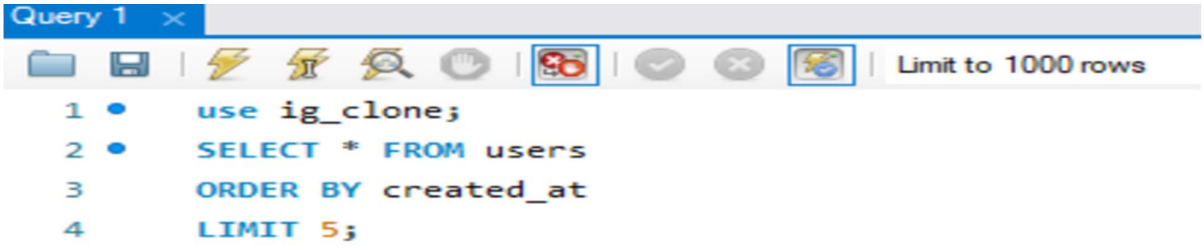
Submission by Swet Parekh

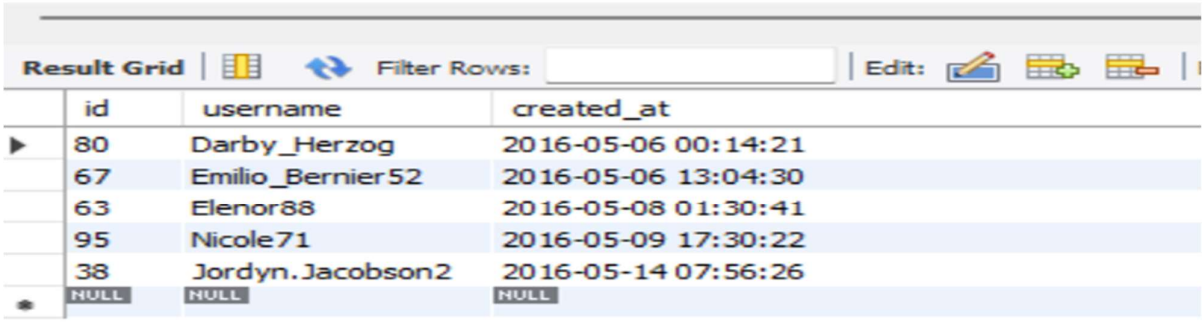
Description:

User analysis is the process by which we track how users engage and interact with our digital product (software or mobile application) in an attempt to derive business insights for marketing, product & development teams.
These insights are then used by teams across the business to launch a new marketing campaign, decide on features to build for an app, track the success of the app by measuring user engagement and improve the experience altogether while helping the business grow.
You are working with the product team of Instagram and the product manager has asked you to provide insights on the questions asked by the management team.

A) Marketing: The marketing team wants to launch some campaigns, and they need your help with the following

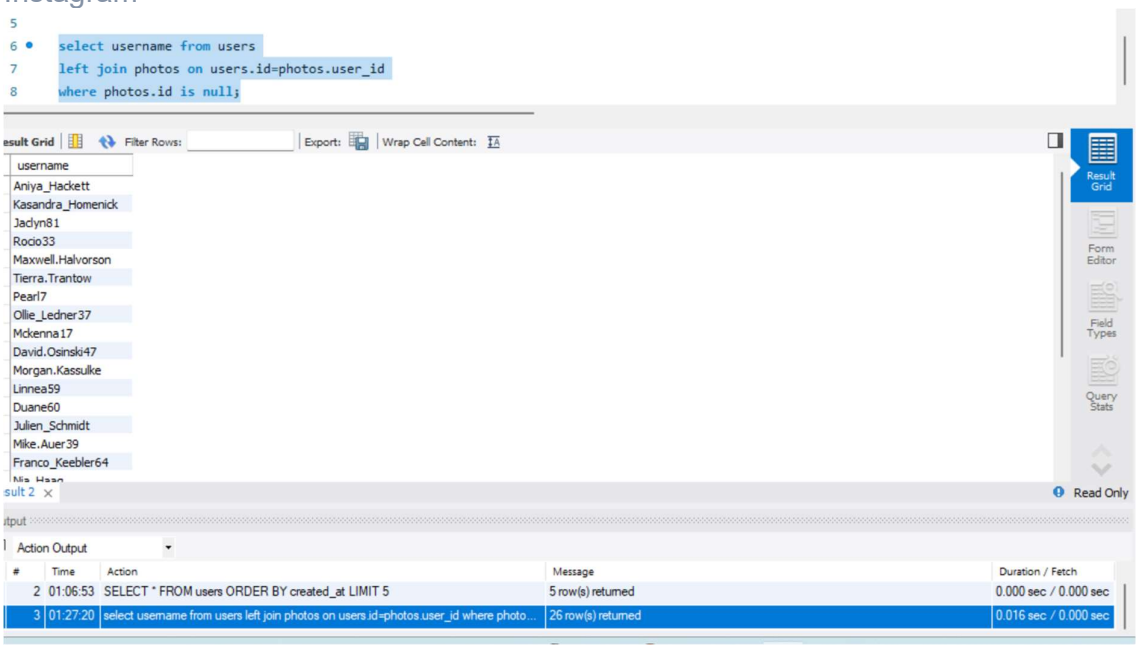
1. **Rewarding Most Loyal Users:** People who have been using the platform for the longest time.
   Your Task: Find the 5 oldest users of the Instagram from the database provided





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

2. **Remind Inactive Users to Start Posting:** By sending them promotional emails to post their 1st photo.
   Your Task: Find the users who have never posted a single photo on Instagram



3. **Declaring Contest Winner:** The team started a contest and the user who gets the most likes on a single photo will win the contest now they wish to declare the winner.
   Your Task: Identify the winner of the contest and provide their details to the team

```
    b.username AS UserName
FROM
    (
        SELECT
            l.*,
            p.user_id
        FROM
            (
                SELECT
                    photo_id,
                    COUNT(*) AS Likes
                FROM
                    ig_clone.likes
                GROUP BY
                    photo_id
                ORDER BY
                    Likes DESC
                LIMIT 1
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

photo_id	Likes	user_id	UserName
145	48	52	Zack_Kemmer93

4. **Hashtag Researching:** A partner brand wants to know, which hashtags to use in the post to reach the most people on the platform.
   Your Task: Identify and suggest the top 5 most commonly used hashtags on the platform

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

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

tag_name	hashtag_count
smile	59
beach	42
party	39
fun	38
concert	24

Result 4

5. **Launch AD Campaign:** The team wants to know, which day would be the best day to launch ADs.
   Your Task: What day of the week do most users register on? Provide insights on when to schedule an ad campaign

```
use ig_clone;
SELECT
    DAYNAME(created_at) AS registration_day,
    COUNT(*) AS user_count
FROM
    ig_clone.users
GROUP BY
    registration_day
ORDER BY
    user_count desc

LIMIT 1;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Cor

registration_day	user_count
Thursday	16

B) Investor Metrics: Our investors want to know if Instagram is performing well and is not becoming redundant like Facebook, they want to assess the app on the following grounds

1. **User Engagement:** Are users still as active and post on Instagram or they are making fewer posts
   Your Task: Provide how many times does average user posts on Instagram. Also, provide the total number of photos on Instagram/total number of users

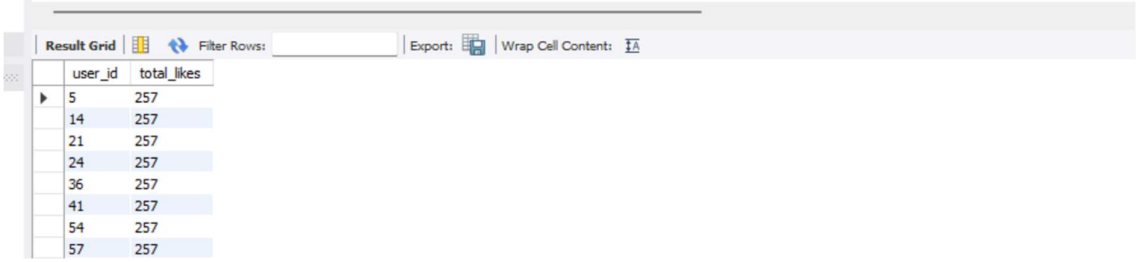
```
SELECT
    COUNT(*) / (SELECT COUNT(*) FROM ig_clone.users) AS average_posts_per_user,
    COUNT(*) AS total_photos,
    (SELECT COUNT(*) FROM ig_clone.users) AS total_users
FROM
    ig_clone.photos;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

average_posts_per_user	total_photos	total_users
2.5700	257	100

2. **Bots & Fake Accounts:** The investors want to know if the platform is crowded with fake and dummy accounts
   Your Task: Provide data on users (bots) who have liked every single photo on the site (since any normal user would not be able to do this).

```sql
SELECT
    user_id,
    COUNT(*) AS total_likes
FROM
    ig_clone.likes
GROUP BY
    user_id
HAVING
    total_likes = (SELECT COUNT(DISTINCT id) FROM ig_clone.photos);
```



Result Grid | Filter Rows: | Export: | Wrap Cell Content:

user_id	total_likes
5	257
14	257
21	257
24	257
36	257
41	257
54	257
57	257

Approach: For this project, I utilized a combination of SQL queries and data analysis techniques to gain insights into Instagram's user engagement and the presence of bots or fake accounts. I extracted relevant data from the database using SQL queries to answer specific questions posed by the investors. The data was then analyzed to provide meaningful insights regarding user activity, average posts per user, the total number of photos, and the identification of users who have liked every single photo on the site.

Tech-Stack Used:

I used SQL Workbench as the software to interact with the database and execute the SQL queries. SQL Workbench is a popular SQL client tool that provides a user-friendly interface for managing and querying databases. It allows for efficient data retrieval and analysis by providing a comprehensive set of features and functionalities. The queries were executed in SQL Workbench to extract the required data and derive insights from the database.

Insights: Through the project, I gained insights into various aspects of Instagram's performance. By calculating the average posts per user, I could assess user engagement and activity on the platform. Furthermore, by determining the total number of photos and users, I obtained a sense of the platform's overall size and reach. Additionally, identifying users who have liked every single photo on the site helped to highlight potential instances of fake or bot accounts, which is crucial for evaluating platform authenticity.

Result:

By utilizing SQL Workbench and executing the SQL queries, I successfully obtained insights into user engagement and the presence of bots or fake accounts on Instagram. The average posts per user metric provided an understanding of user activity, while identifying users who liked every photo helped to identify potential fake accounts. These insights enabled me to address investor concerns about the platform's performance and authenticity. The project's outcome was a comprehensive understanding of Instagram's user dynamics, which proved valuable for decision-making and assessing the platform's overall health.