Himanshu Jani - Trianity - Project - 2 - Instragram User Analytics

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

```
-- Task: Find the 5 oldest users of Instagram

-- Purpose: Rewarding Most Loyal Users - People who have been using the platform for the longest time.

-- Step 1: Select the necessary columns from the users table

SELECT id, username, created_at
FROM ig_clone.users

-- Step 2: Order the results based on the date they joined the platform (created_at) in ascending order

ORDER BY created_at

-- Step 3: Limit the results to the top 5 oldest users

LIMIT 5;
```

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

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

-- Step 1: Select the necessary columns from the users table

SELECT u.id, u.username
FROM ig_clone.users AS u

-- Step 2: Left join the users table with the photos table to find users who have never posted a photo

LEFT JOIN ig_clone.photos AS p ON u.id = p.user_id

-- Step 3: Filter out the users who have never posted a photo

WHERE p.id IS NULL;
```

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

```
-- Task: Identify the winner of the contest based on the most likes on a single photo

-- Step 1: Count the number of likes for each photo and select the maximum count

SELECT photo_id, COUNT(*) AS like_count
FROM ig_clone.likes
GROUP BY photo_id
ORDER BY like_count DESC
LIMIT 1;

-- Step 2: Retrieve the details of the user who owns the winning photo


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

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

```
-- Task: Identify the top 5 most commonly used hashtags on the platform

-- Step 1: Count the occurrences of each hashtag in the photo_tags table

SELECT tag_id, COUNT(*) AS hashtag_count
FROM ig_clone.photo_tags
GROUP BY tag_id
ORDER BY hashtag_count DESC
LIMIT 5;

-- Step 2: Retrieve the details of the top 5 hashtags

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

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

```
-- Task: Determine the day of the week when most users register on Instagram

-- Purpose: Provide insights on when to schedule an ad campaign

-- Step 1: Extract the day of the week from the registration date

SELECT DAYNAME(created_at) AS registration_day, COUNT(*) AS registration_count
FROM ig_clone.users
GROUP BY registration_day
ORDER BY registration_count DESC;
```

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

```
-- User Engagement Metrics

-- Average number of posts per user


-- Total number of photos on Instagram / Total number of users

SELECT

   (SELECT AVG(post_count) AS average_posts_per_user
   FROM (
      SELECT COUNT(*) AS post_count
      FROM ig_clone.photos
      GROUP BY user_id
   ) AS subquery) AS average_posts_per_user,

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

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).

```
-- Task: Identify users who have liked every single photo on the site (potentially bots)

-- Step 1: Count the total number of photos
SELECT COUNT(*) AS total_photos
FROM ig_clone.photos;

-- Step 2: Retrieve the users who have liked every photo

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

Project Description

Give a brief about your project description i.e. what is this project about, how are you going to handle the things and what are the things that you are going to find out through the project.

Project Description: Instagram User Analytics

The Instagram User Analytics project aims to analyze various aspects of user behavior and engagement on the Instagram platform. Using the provided database schema, we will use SQL queries to extract meaningful information and metrics that can help in understanding the platform's performance and user choices.

I have performed analysis step by step and divided it into 3 parts.

Database Schema: As we have a well-defined database schema that includes tables for users, photos, comments, likes, follows, tags, and their relationships.

SQL Queries: I have utilized SQL queries to extract relevant information from the database, allowing us to answer specific questions and generate required reports as stakeholders requirements.

Data Analysis: By analyzing the data, I can derive key metrics and patterns related to user engagement, posting frequency, popularity of hashtags, user registration trends, and potential presence of bots or fake accounts.

Seeking Key Information which can give us insights:

- User Engagement:
- Popularity of Hashtags:
- User Registration Trends:
- Bot Detection:

These key information metrics can be used by investors, brands, and Instagram's management to make data-driven decisions, improve user experience, and improve the overall platform effectiveness.

Approach

Write a short paragraph about your approach towards the project and how you have executed it.

For the Instagram User Analytics project, I have followed a systematic approach to extract

insights from the provided database schema.

Understanding the Schema:

First and foremost understanding of database schema is utmost important to begin with I have carefully examined the database schema to understand table and their relationship to extract relevant information.

Formulating Queries:

Based on the given instruction of project requirements , I have formulated SQL queries to retrieve required information and ensured that queries are giving correct information and runs error free and give detailed comments What I have done and logic behind it.

Tech-Stack Used

To execute the Instagram User Analytics project, I have utilized the **DB-fiddle option** for storing the database and running queries as it allows us to run SQL queries directly within the browser.

Insights

Jot down the insights and the knowledge you gained while making the project. You need to write that what do you infer about the things. Make sure its brief and up to the point only. For Eg. If you got a graph then what do you understand by the graph, what changes can you make or what can you derive from the graph.

Here are the **4 Key Info Metrics** I can say we have achieved while doing this project.

1. **User Engagement:** The average number of posts per user on Instagram indicates the level of user activity and posting frequency. By monitoring this metric we can identify any changes in user behavior .

2. **Popularity of Hashtags:** Analyzing the most commonly used hashtags on Instagram helps in understanding the trending topics and interests of users. By incorporating relevant and popular hashtags in their content, they can maximize visibility and connect with a broader audience.

3. **User Registration Trends:** Identifying the day of the week when most users register on Instagram allows for effective planning of ad campaigns.

4. **Bot Detection:** Users who have liked every single photo on the site may indicate the presence of bots or fake accounts. It also assists in identifying potential issues related to spam or automated activity.

Result

Mention what have you achieved while making the project and how do you think it has helped you.

During this entire project I have gained insights into user behavior, engagement patterns, and potential areas of concern such as fake accounts or bots. These insights can guide decision-making processes, help optimize strategies.

This also on personal level took me to the depth of my SQL knowledge and it has tested it thoroughly and It has also developed my analytical abilities and given me glimpse how real world scenario unfold when we start performing analytics so many times I stuck while completing this project but eventually I was able to pulled it surely I can say that I have improved my proficiency in SQL.