Project Description:

In this project, we will use Instagram user data to analyze user interactions and their engagement with the Instagram app to provide valuable insights that can help the business grow by taking data driven decisions. User analysis involves tracking how users engage with a digital product, such as a software application or mobile app. The insights derived from the analysis can be used by various teams within the business. For example, the marketing team might use these insights to launch a new campaign, the product team might use them to decide on new features to build, and the development team might use them to improve the overall user experience.

We will use MySQL Workbench as the tool to analyze the Instagram user data and to answer the questions posed by the management team.

Tech-Stack Used:

For this project, we used MySQL Workbench 8.0 to analyze the user data provided and get the insights from it. We used this tool as it is open-source, user-friendly, easy to install, it is secure and most widely used database management system in the world.

Approach:

The dataset provided was first loaded in MySQL Workbench and database was created. Then, the database was selected with **use** and checked the tables present using **select** inside the data. Checked for any missing values with **describe.** Then, we used types of **joins** (**inner, left**) and **with** to join the tables to fetch the required records to answer the questions posed by the management team.

A) Marketing Analysis:

1. Loyal User Reward:

Instagram*

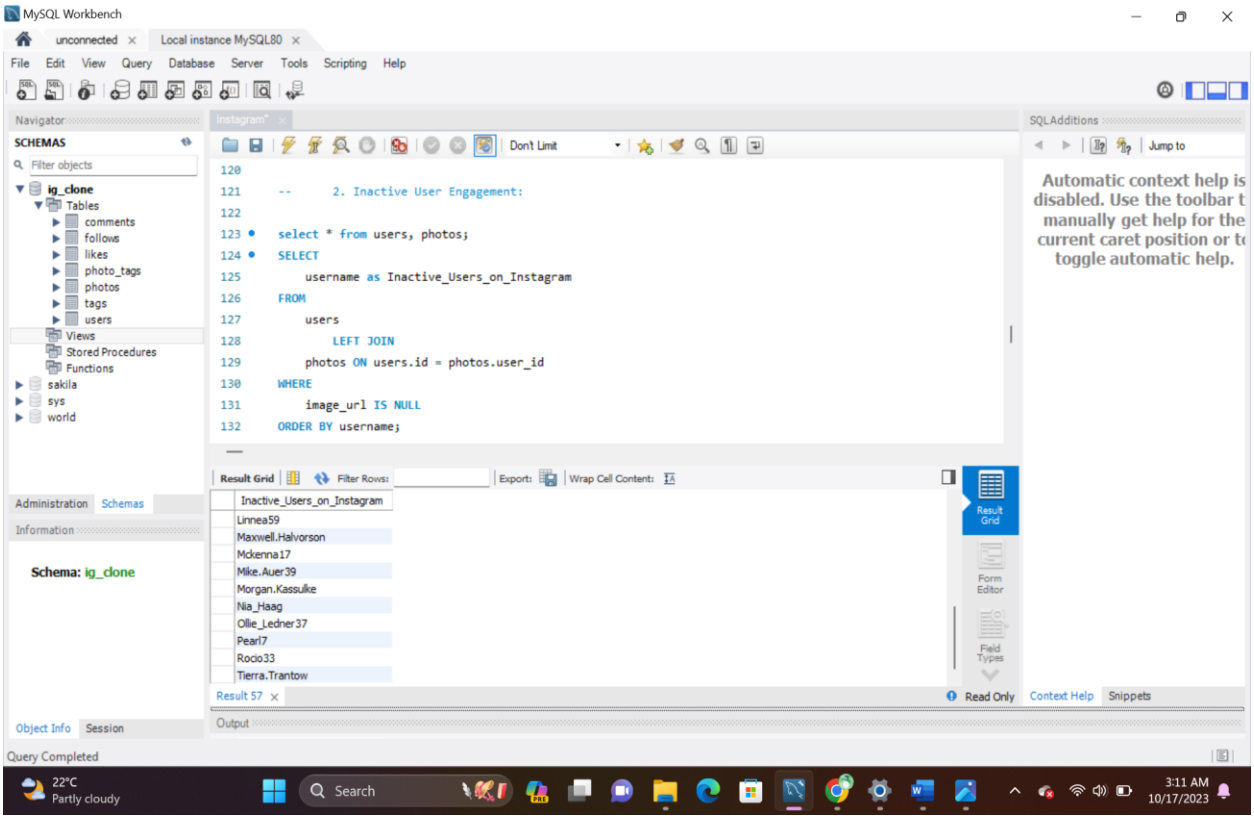
```sql
-- A) Marketing Analysis:

--      1. Loyal User Reward:
SELECT
    username, created_at
FROM
    users
ORDER BY created_at
LIMIT 5;
```

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

username	created_at
Darby_Herzog	2016-05-06 00:14:21
Emilio_Bernier52	2016-05-06 13:04:30
Elenor88	2016-05-08 01:30:41
Nicole71	2016-05-09 17:30:22
Jordyn.Jacobson2	2016-05-14 07:56:26

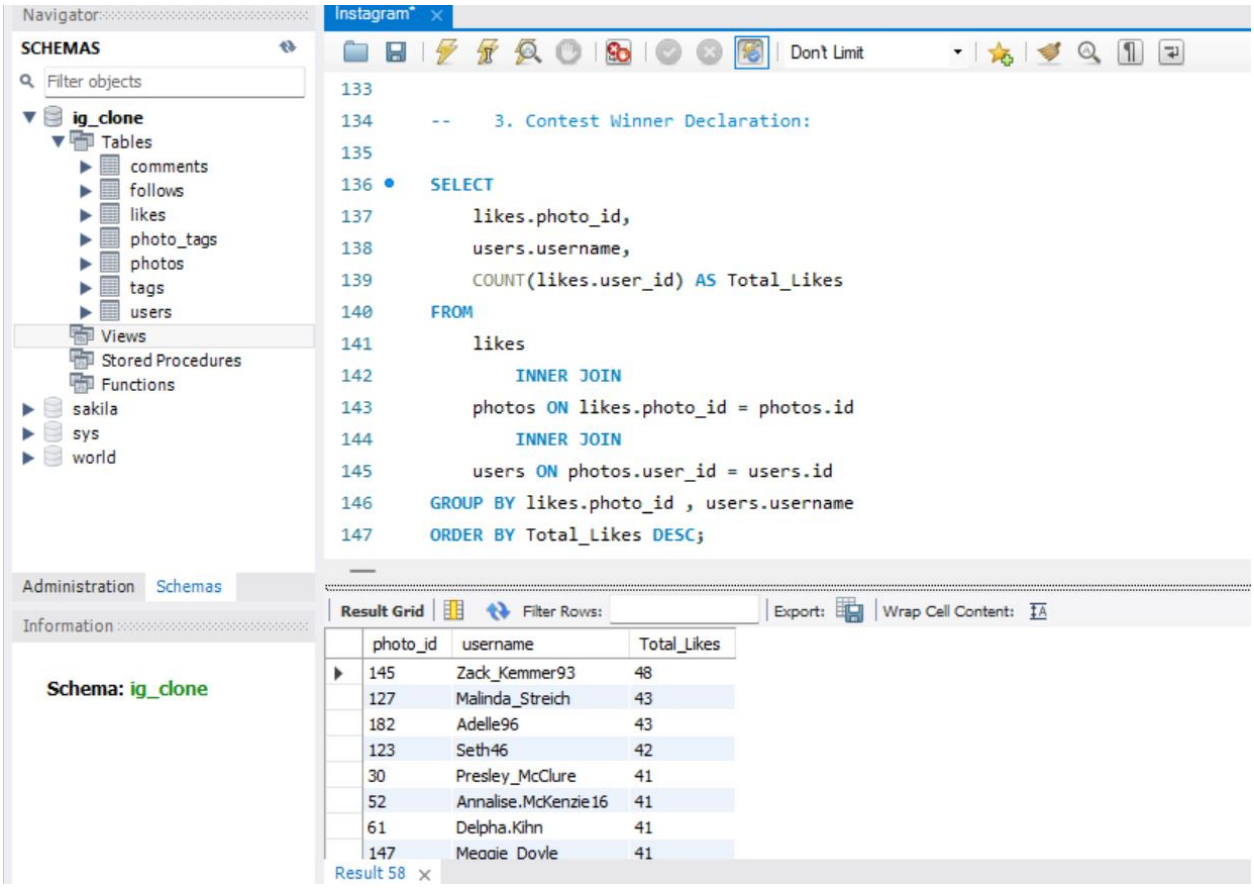
Here, we ordered the data with respect to **created_at** to find out top 5 oldest users on Instagram and found that the first user that registered was on 2016-05-06 00:14:21.

2. Inactive User Engagement:



In this, we joined users and photos table with **left join** to find out the users who have never posted a single photo on Instagram. We used to filter the records here based on **where image_url is null** to get the desired result.

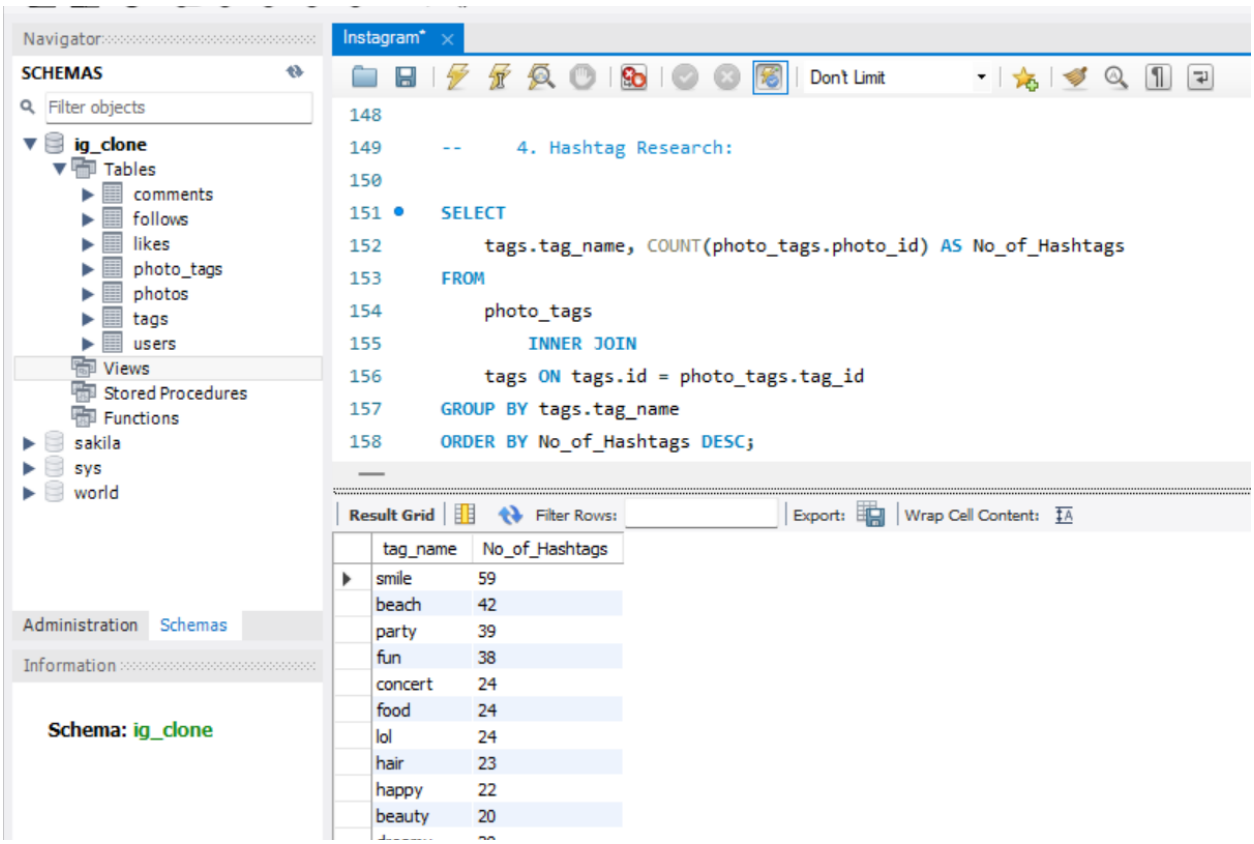
3. Contest Winner Declaration:



In this part, we needed to find out the user who has the most number of likes on a single photo. So, we joined three tables here which are likes, photos, and users using **inner join**. We counted the number of likes by counting the **user_id** in likes table .

We found out that the user with the most likes had a total of **48 likes**.

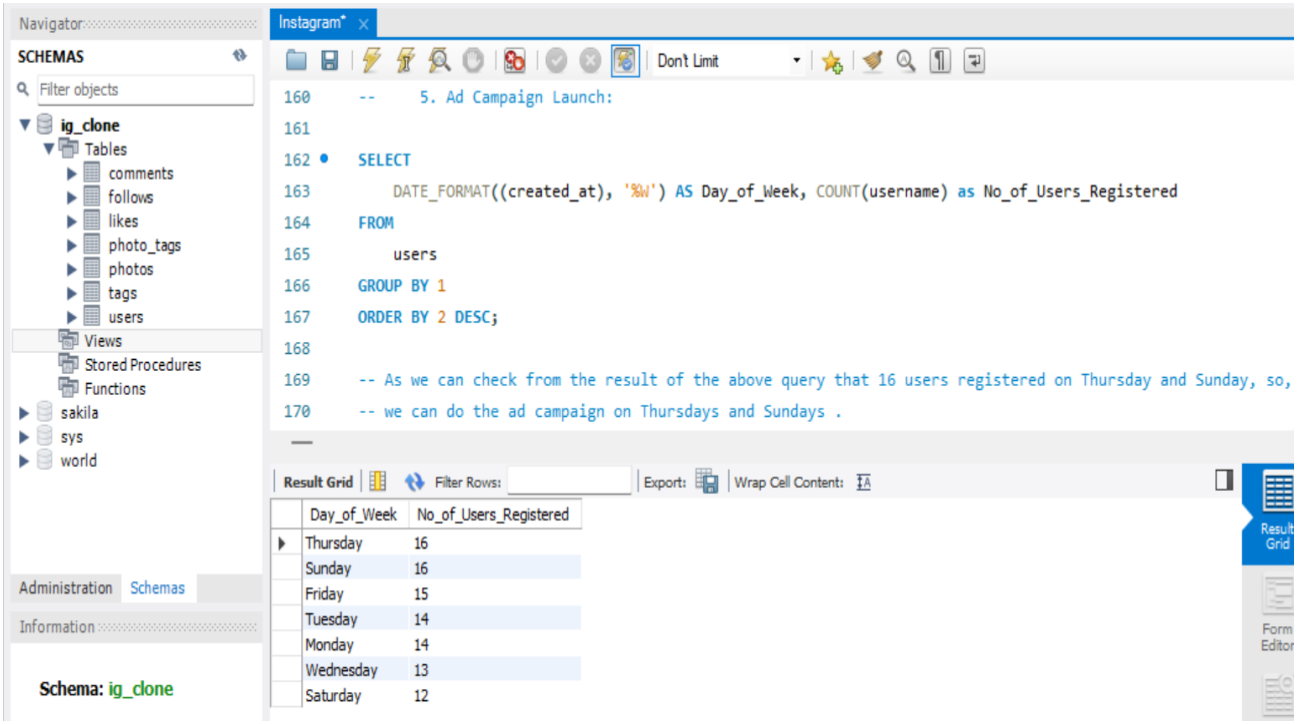
4. Hashtag Research:



Here, we found out the most popular hashtags that could be used in posts to reach to the most people. To find out the popular hashtags, we counted the number of times a particular hashtag has been used in a post. For this, we used **photo_tags** and **tags** table and joined using **inner join**. Then, the records were grouped by the **tag_name** and ordered by number of times it has been used in descending order.

We found out that **smile, beach, party, fun, concert,** are some of the most popular and common hashtags that people use in their posts.

5. Ad Campaign Launch:

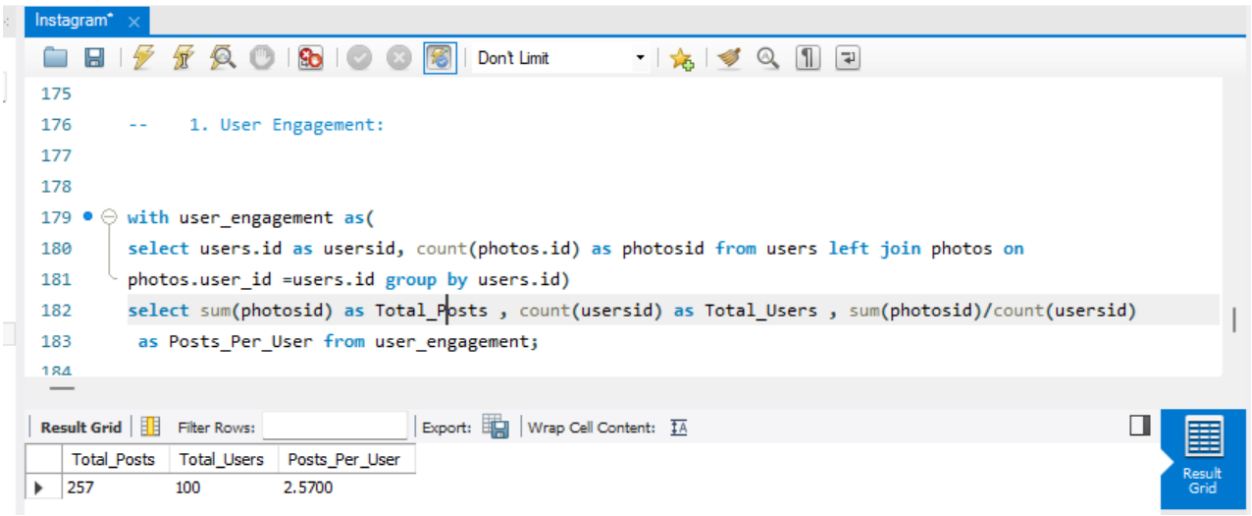


Here, we wanted to find out the best day of the week to launch a Ad campaign. To get the result, we used **DATE_FORMAT()** and fetched the day from the **created_at** column in the users table. Then, we counted the total number of users that were registered on each day of week and arranged the result in descending order in order to find out on which day most users register themselves on Instagram.

Based upon the finding, we see that on **Thursday** and **Sunday** most users registered themselves on Instagram. So, a ad-campaign should be organized on these days.

B) Investor Metrics:

1. User Engagement:

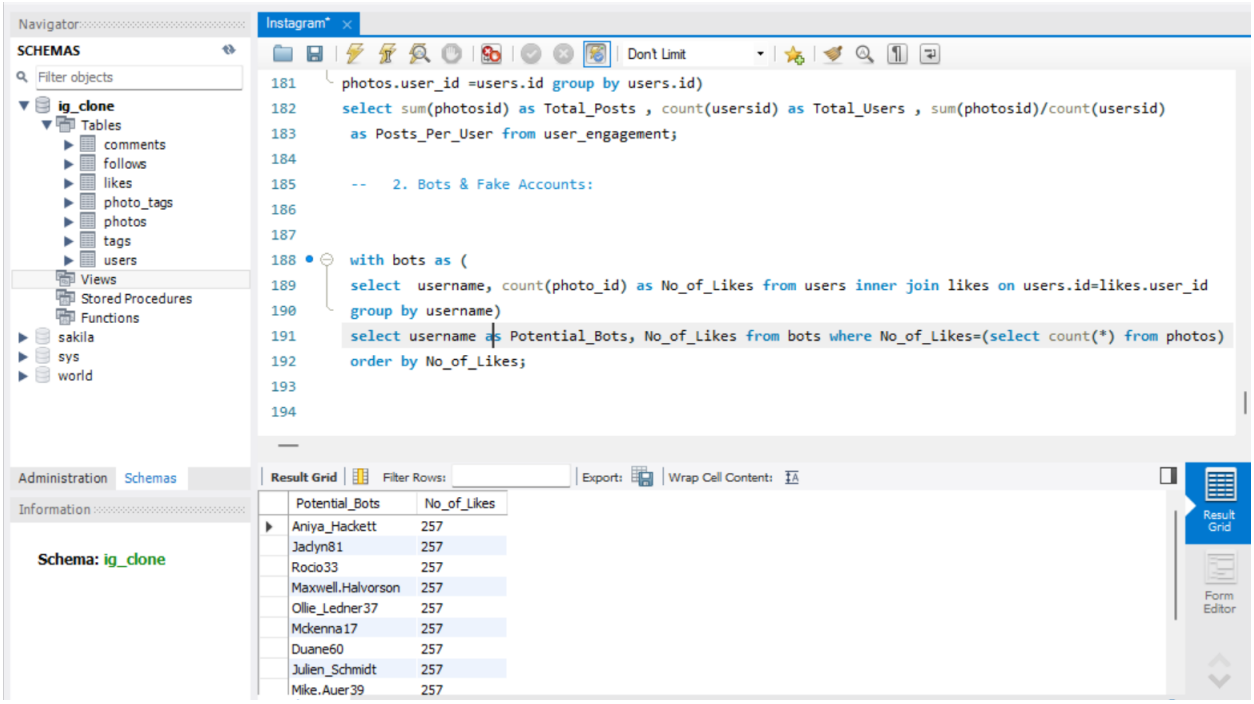


Here, the investors wanted to find out if the users are still active and posting on Instagram or if they are making fewer posts. So, we found out the average number of posts per user i.e., on an average how many posts a user is doing. To get this result, we needed to calculate the total users on Instagram and the total number of posts that have been made.

We used **with** and joined tables users and photos using **left join** and to calculate the total posts we used **sum()** and for total users we used **count().** To find out the average no of posts per user we divided the total posts by total users.

Based on the result, we see that on an average a user makes around **2.7** posts or we can round off and say **3** posts where the total users are **100** and total posts done are **257.**

2. Bots & Fake Accounts:



Here, we wanted to find out if there are any users registered on Instagram as bots which are fake and dummy accounts. In order to achieve this, we checked if there are any users who have liked every single photo on the website as this is typically not possible for a normal user.

To get the desired result, we joined the tables **users** and **likes** using **inner join** and counted the total number of likes from **count()** on **photo_id.** Then, we counted the total number of records present in **photos** table and filtered our records as the **username** of the user whose total number of likes is equal to the total number of photos.

Analyzing this, we found out there are **13** such users registered which could be a potential bot or a fake and dummy account.

Insights:

- We found out that the first user that registered on Instagram was on **2016-05-06 00:14:21**.
- We found out that there is a total of **26** such users who have never posted a single photo.
- We found out that the most number of likes on a single photo for a user is **48.**
- We found out that the most common/popular hashtags used on posts by the users are **smile, beach, party, fun, concert.**
- We found out that on **Thursdays** and **Sundays**, maximum number of users are registered on the website.
- We found out that on an average a user posts around **3 photos.**
- We found out that there is a total of **13** such users who could be a possible bot/fake/dummy account.

Result:

Based upon the findings, we are able to detect that we need more engagement from the users on the platform as the number of posts per user seems to be less. In order to achieve this, we can start a ad-campaign on the days that most users are registered. Also, we found out that out of a total of 100 users, 13 of them are not real users. So, we should either deactivate them or does not count them as a user for metrics.

We saw the most popular hashtags that people use on their posts, so, we can use them to reach a larger group of people while doing any promotions. Also, there are 26 users who have never posted a single photo on Instagram, so, we can offer these users with some sort of promotion to increase the engagement or target them with some specific ads. Also, we checked the oldest users on the app, so, we can do promotions by giving them some sort of benefits or services for free for a period of time.