*Instagram user analytics*

# Project Description

This project aims to examine how users interact with and engage on the Instagram app, yielding insights that are crucial for business expansion. It entails monitoring user engagement with digital products like software or mobile applications. These insights are instrumental for various business teams, aiding in the initiation of new campaigns or informing decisions regarding the development of new features. Additionally, the development team can utilize this information to enhance the user experience.

# Approach

To perform the required tasks and complete the project , I have used SQL queries using MySQL

Commandline client . According to the instructions provided to create the database and the

corresponding tables , I feeded the data into MySQL and executed the appropriate queries to

get the required insights.

Values were generated and entered into the database by executing the DDL (Data Definition Language) and DML (Data Manipulation Language) SQL queries supplied by the product manager, in accordance with the project requirements, utilizing the MySQL Workbench for the MySQL database.

The process entailed the following steps:

1. Database creation was initiated.
2. A ‘users’ table was established to catalog user details, assigning each a distinct ID, username, and creation timestamp—a standard structure for user data management in databases.
3. A ‘photos’ table was set up to record photo details, with each photo assigned a unique ID, an image URL, the uploader’s user ID, and an upload timestamp.
4. A ‘comments’ table was constructed to keep track of user comments on photos, with each comment having a unique ID, comment text, the commenter’s user ID, the associated photo’s ID, and a creation timestamp.
5. A ‘likes’ table was created to document likes, linking each like to a user and a photo, accompanied by the timestamp of the like.
6. A ‘followers’ table was formed to store data on user followings, capturing the follower’s and followee’s IDs, along with the timestamp of the following action.
7. A ‘tags’ table was developed to store tag information, with each tag given a unique ID, a tag name, and a creation timestamp.
8. A ‘photo\_tags’ junction table was created to facilitate the many-to-many association between photos and tags, with each row indicating a link between a specific photo and tag.
9. Values were populated into all tables using the provided datasets.
10. Upon database completion, necessary insights were extracted from the database tables through SQL queries executed in MySQL Workbench.

To perform the required tasks and complete the project , I have used SQL queries using MySQL

Commandline client . According to the instructions provided to create the database and the

corresponding tables , I feeded the data into MySQL and executed the appropriate queries to

get the required insights.

To perform the required tasks and complete the project , I have used SQL queries using MySQL

Commandline client . According to the instructions provided to create the database and the

corresponding tables , I feeded the data into MySQL and executed the appropriate queries to

get the required insights.

To perform the required tasks and complete the project , I have used SQL queries using MySQL

Commandline client . According to the instructions provided to create the database and the

corresponding tables , I feeded the data into MySQL and executed the appropriate queries to

get the required insights.

To perform the required tasks and complete the project , I have used SQL queries using MySQL

Commandline client . According to the instructions provided to create the database and the

corresponding tables , I feeded the data into MySQL and executed the appropriate queries to

get the required insights.

# Tech-Stack Used

* **MySQL Workbench**
* MySQL Community Server — GPL Version 8.0.29
* **SQL Server Management Studio (SSMS)**

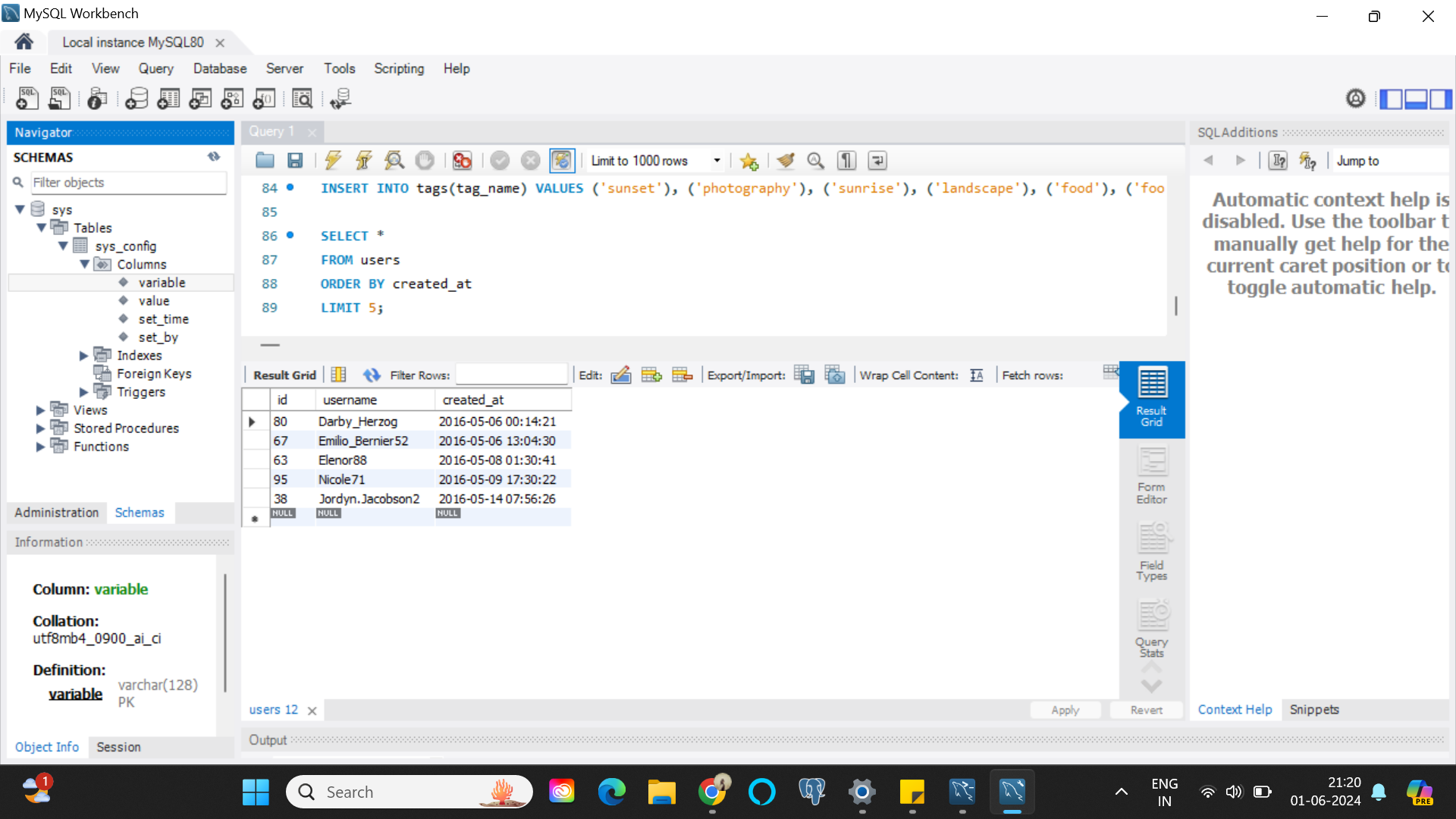
All software & tools are chosen based on their compatibility with the respective database systems, the specific needs of the project, and the preference for comprehensive tools that facilitate efficient database management and development. MySQL Community Server — GPL Version 8.0.29 is a free and open-source relational database management system that uses SQL.MySQL Workbench is typically chosen for MySQL database projects, while SSMS is the go-to for projects involving SQL Server databases. Their selection is driven by the goal to leverage their strengths in database design, management, and optimization to ensure the project’s success.

# Insights

A snapshot of both the query as well as the obtained output has been listed for each question posed by the management team. Each insight observed is highlighted in blue colour. Following insights can be inferred from each query:

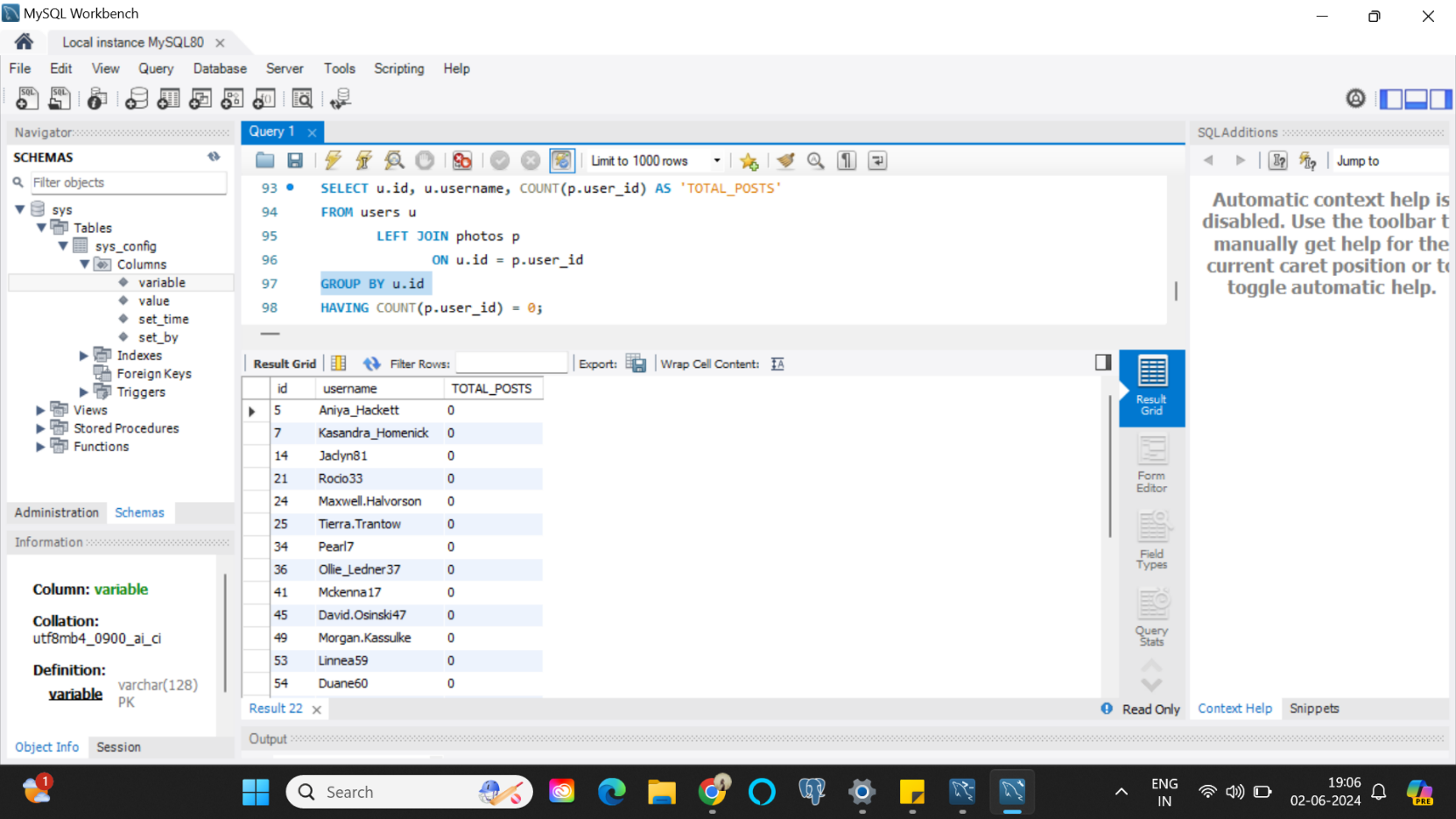
## A) Marketing Analysis:

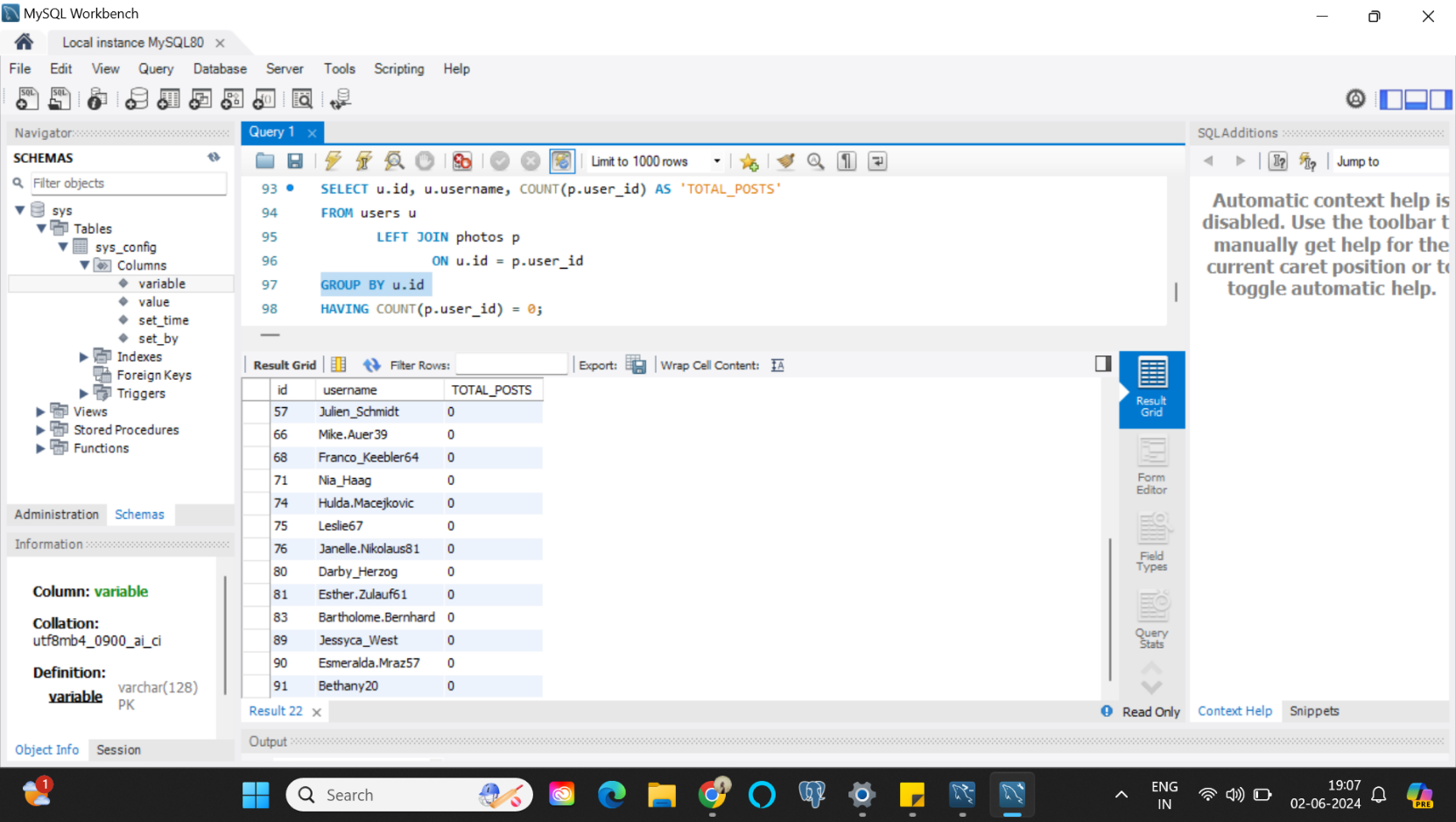
1. Loyal User Reward: Identify the five oldest users on Instagram from the provided database.



The five earliest registered users on Instagram, as per the database, are: Darby\_herzog, Emilio\_Bernier52, Elenor88, Nicole71, and Jordyn.Jacobson2. It is observed that these initial Instagram members created their accounts in May 2016.

1. Inactive User Engagement: Identify users who have never posted a single photo on Instagram.



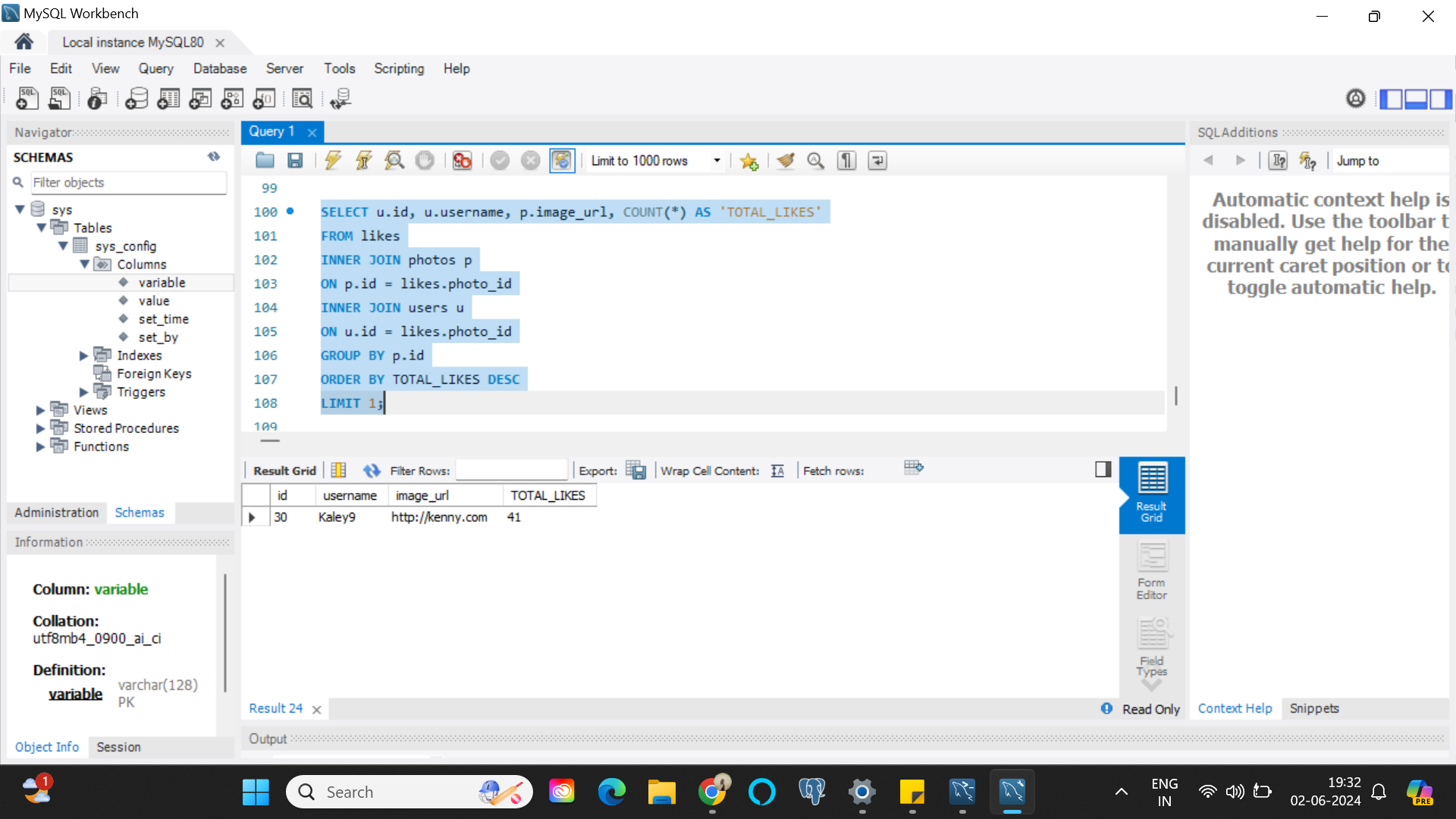


The column labelled ‘username’ lists all the users who have not uploaded any photos on Instagram up to the point of this dataset’s recording.

According to the chart, 26 users have not made any posts on Instagram out of 100 total users, which is 26% of the total number of users using instagram.

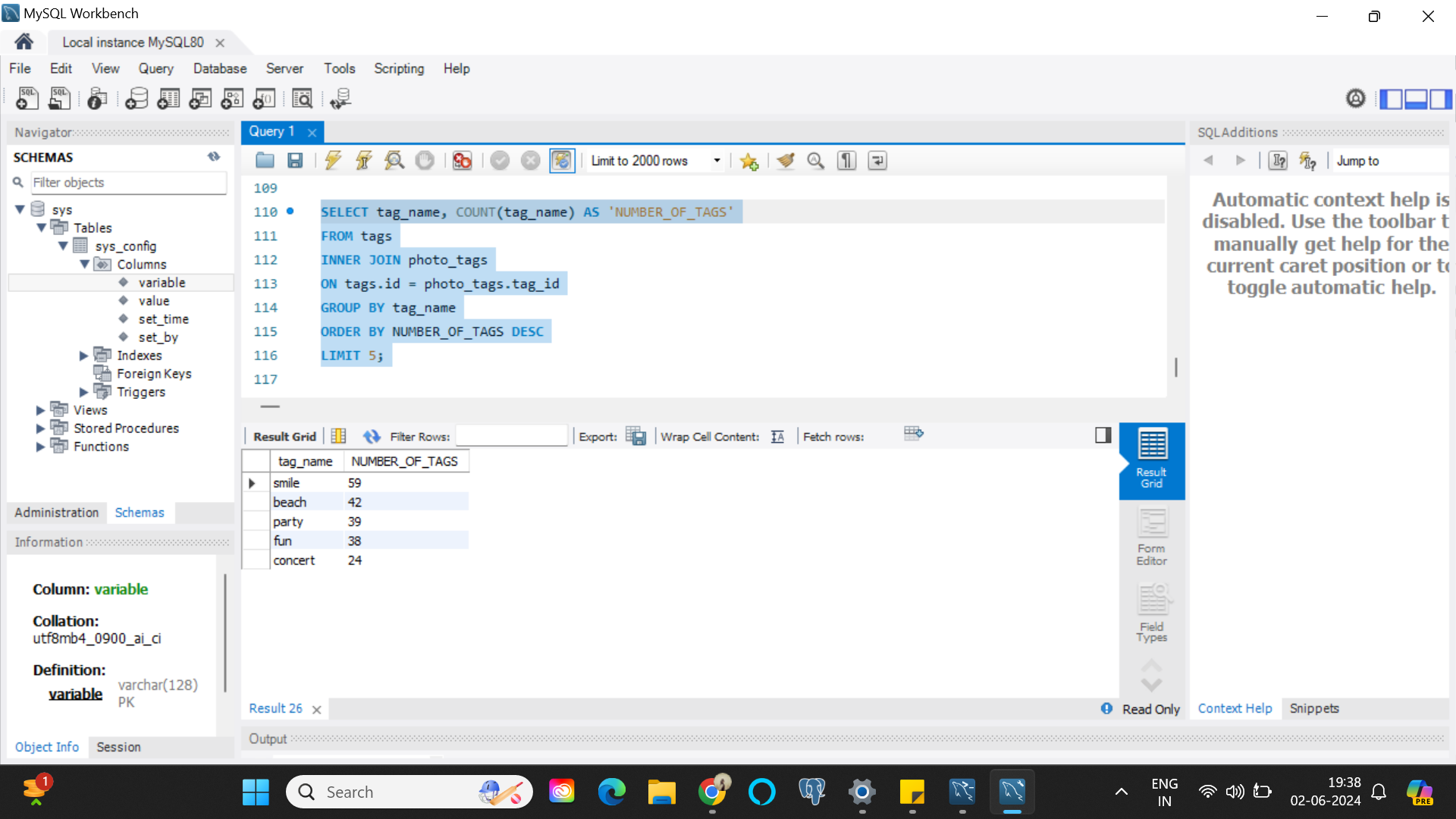
Additionally, the majority of users have shared five or fewer photos, while a minority have shared more than five.

1. Contest Winner Declaration:  Determine the winner of the contest and provide their details to the team.



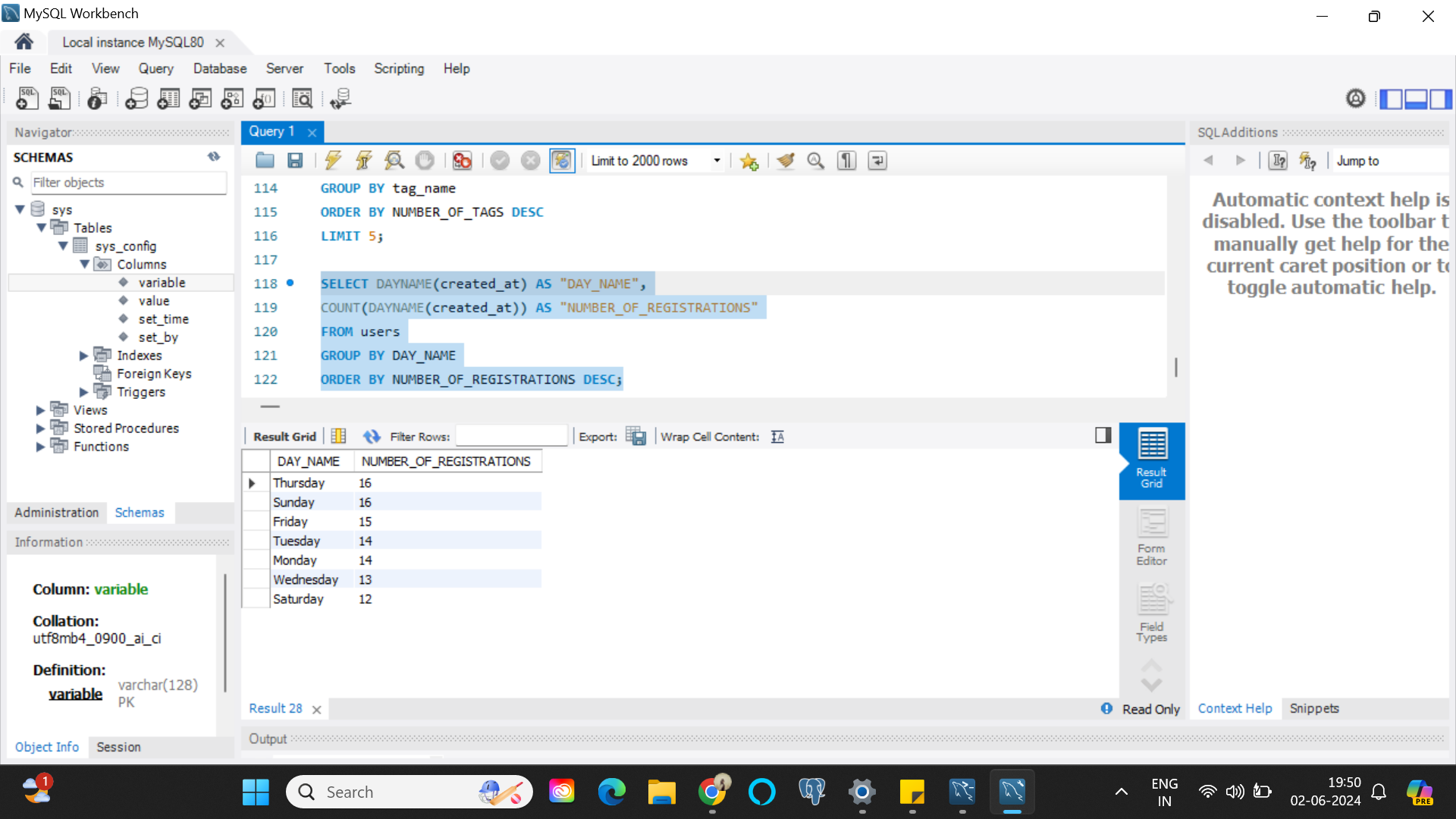
A user by the username ‘Kaley9’ has a total of 41 likes on a single picture and is the winner of the contest.

1. Hashtag Research: Identify and suggest the top five most commonly used hashtags on the platform.



Most popular tags to use to increase reach would be smile, beach, party, fun, concert.

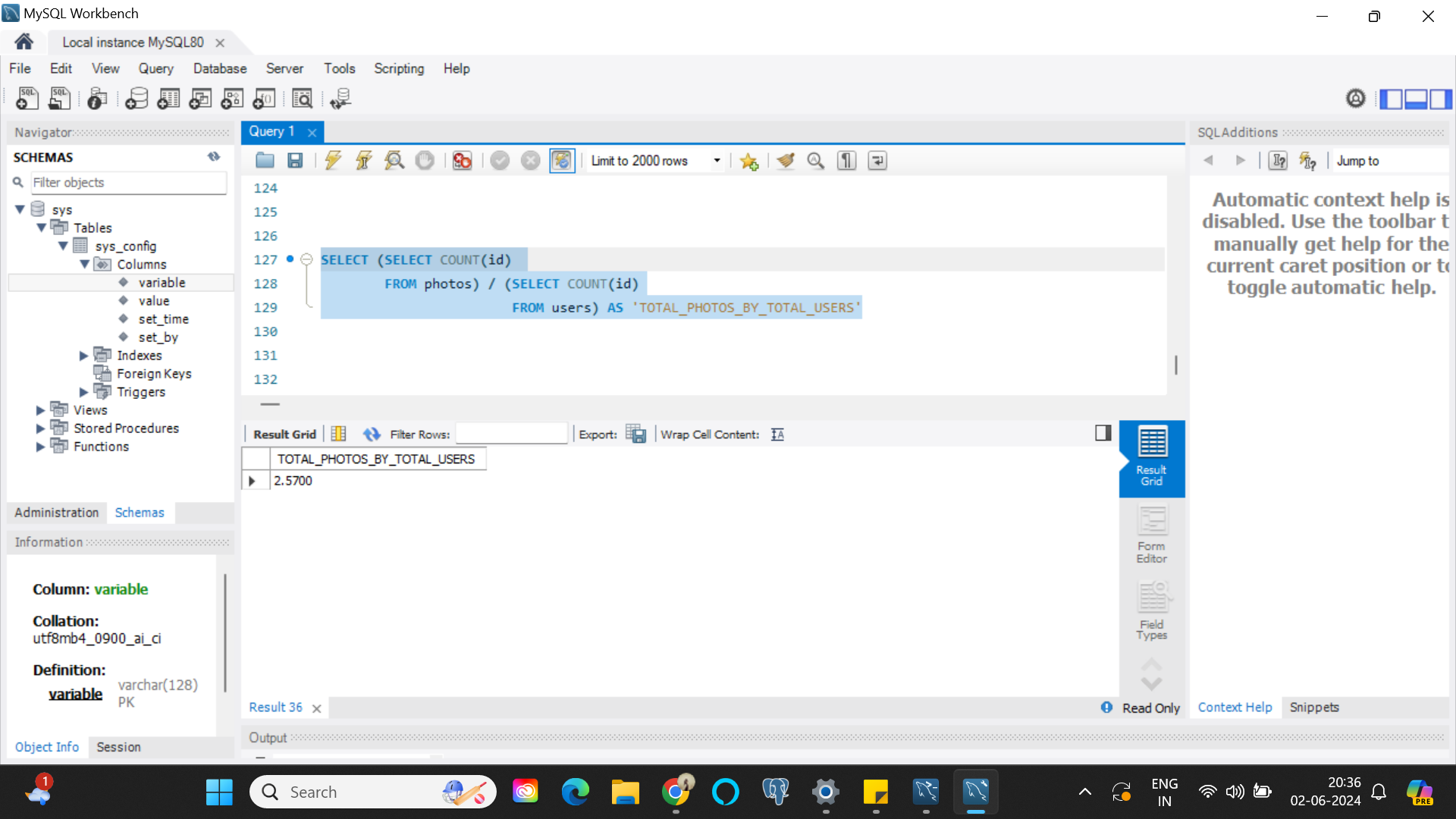
1. Ad Campaign Launch: Determine the day of the week when most users register on Instagram. Provide insights on when to schedule an ad campaign.

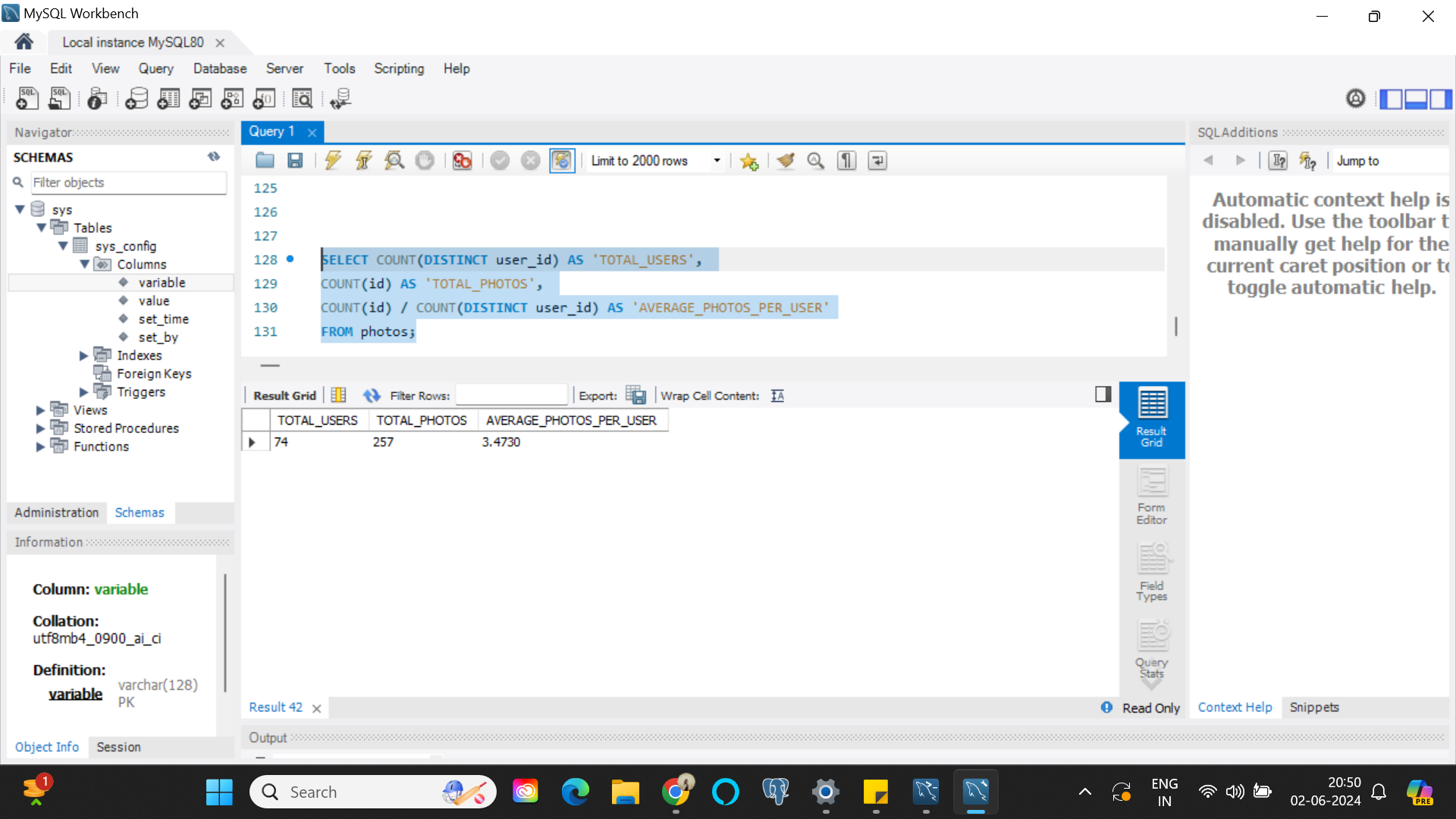


From the above data, we can infer that most users registered on Thursday and Sunday, while Saturday sees the lowest number of registrations. Essentially any Ad campaigns that need to be run should be run on Thursdays and Sundays.

B) Investor Metrics:

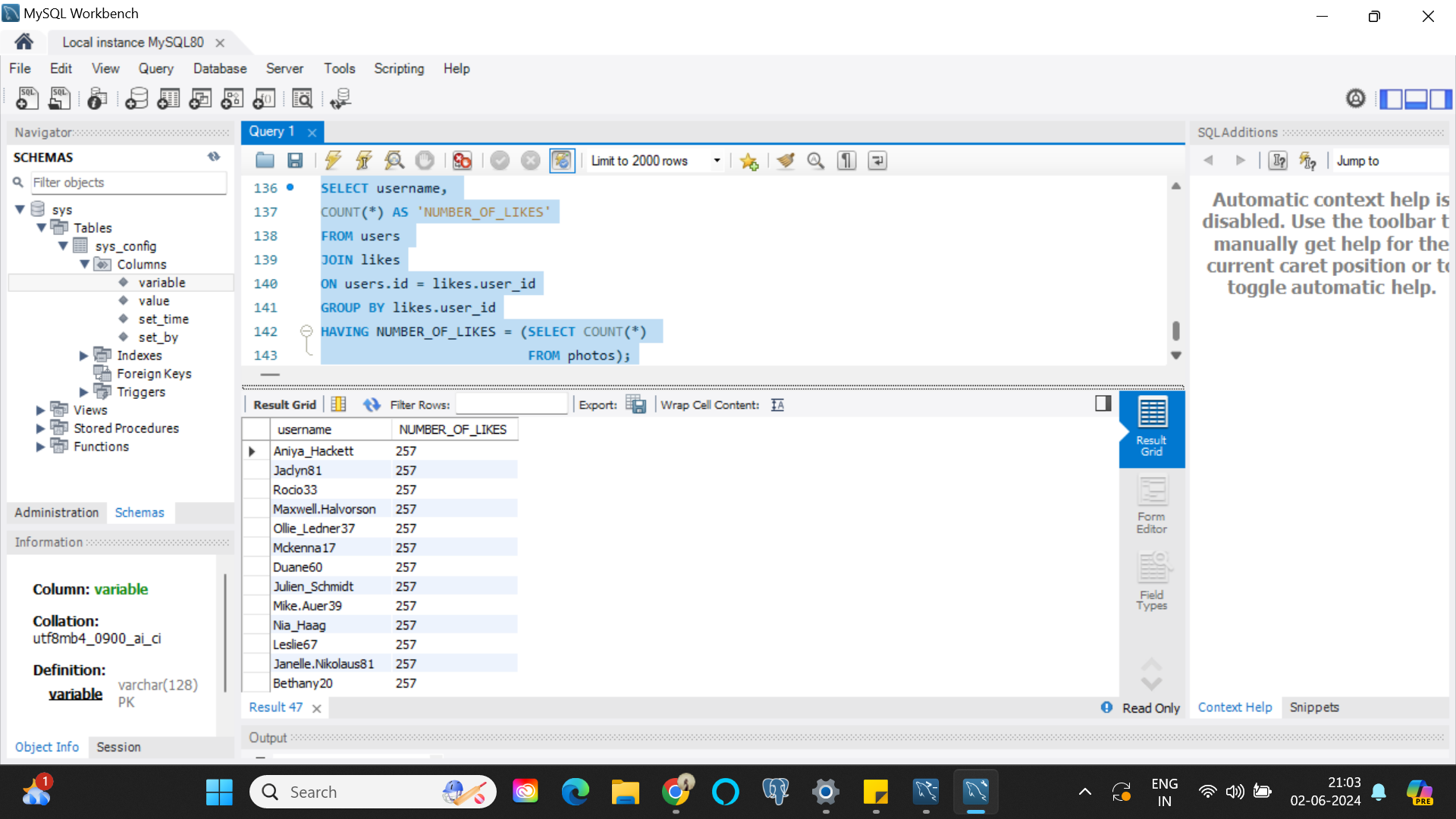
1. User Engagement: Calculate the average number of posts per user on Instagram. Also, provide the total number of photos on Instagram divided by the total number of users.





Here, we can see the average photos per user is 3.4730 and the ratio of total photos to total users is 2.57.

1. Bots & Fake Accounts: Identify users (potential bots) who have liked every single photo on the site, as this is not typically possible for a normal user.



There are a total of 13 potential bots, as can be seen from the above screenshot. 13 out of 100 users in the given data can be considered Bots based on the fact that they have liked each and every picture present on Instagram, which essentially is not possible for a normal user to do.

# Result

This project deepened my comprehension of the critical role data analysis plays in enabling organizations to make informed data-driven decisions. Through this project, which utilized Instagram data, I gained insights into a variety of queries such as identifying the longest-serving users on the platform, pinpointing inactive users, determining the most effective hashtags for promotional content to maximize reach, assessing the prevalence of fake or bot accounts, and evaluating whether the platform is experiencing growth or has plateaued. Consecutively, I have learnt the importance of user engagement and how it can be increased with the help of data analytics