PROJECT 2

# INSTAGRAM USER ANALYTICS

**PROJECT DESCRIPTION:**

THIS PROJECT ANALYZES INSTAGRAM'S DATA THAT TELLS US ABOUT USER BHEAVIOUR, DETERMINING POPULAR HASHTAGS, ANALYZING PREFERABLE DAYS WHEN THE USERS ARE MORE ENGAGED ON THE APP AND THE MAIN GOAL IS PROVIDE MEANINGFUL INSIGHTS THAT CAN GIVE US IDEAS AND STRATEGIES TO GROW THE BUSINESS.

**APPROACH:**

1. Firstly, I understood the data and reviewed all the tables like users, likes, photos etc. and got some idea how the relation can be built.
2. After understanding all the questions, executed the queries, mainly focused on filtering, joining tables ,also used some aggregate functions like count, average etc.
3. Then analyzed the results and interpreted findings and made actionable insights.

**TECH-STACK USED:**

- **SQL SERVER MANAGEMENT STUDIO (SSMS)-**I preferred this software as it provides user friendly interface and seamless experience and supports all the SQL server features including advanced data management.

**QUERIES** -

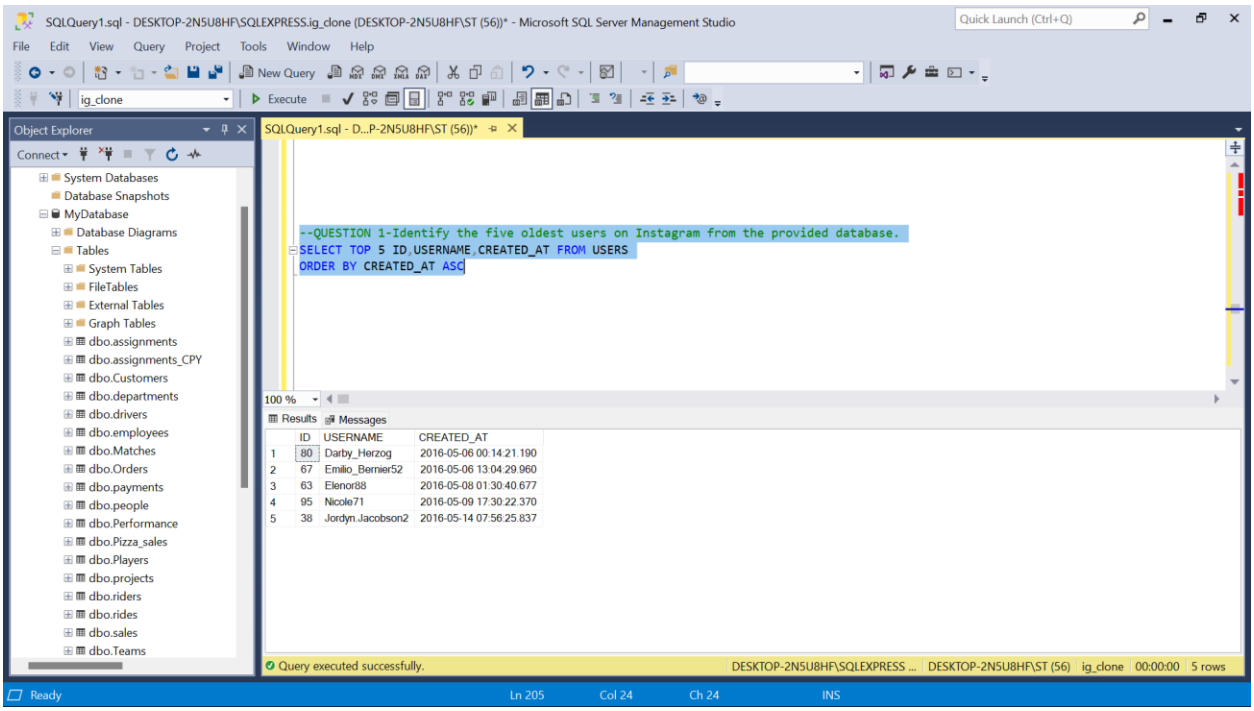
**--QUESTION 1-Identify the five oldest users on Instagram from the provided database.**

```
SELECT TOP 5 ID, USERNAME, CREATED_AT FROM USERS

ORDER BY CREATED_AT ASC
```

**CONCLUSION-**

**THE MOST LOYAL USERS WHO HAVE BEEN USING THE PLATFORM FOR THE LONGEST TIME ARE-**

- **Darby Herzog**
- **Emilio_Bernier52**
- **Elenor88**
- **Nicole71**
- **Jordyn.Jacobson2**




## --QUESTION 2-Identify users who have never posted a single photo on Instagram.

```
SELECT U.ID,U.USERNAME

FROM USERS U

LEFT JOIN

PHOTOS P ON U.ID=P.USER_ID

WHERE P.ID IS NULL
```

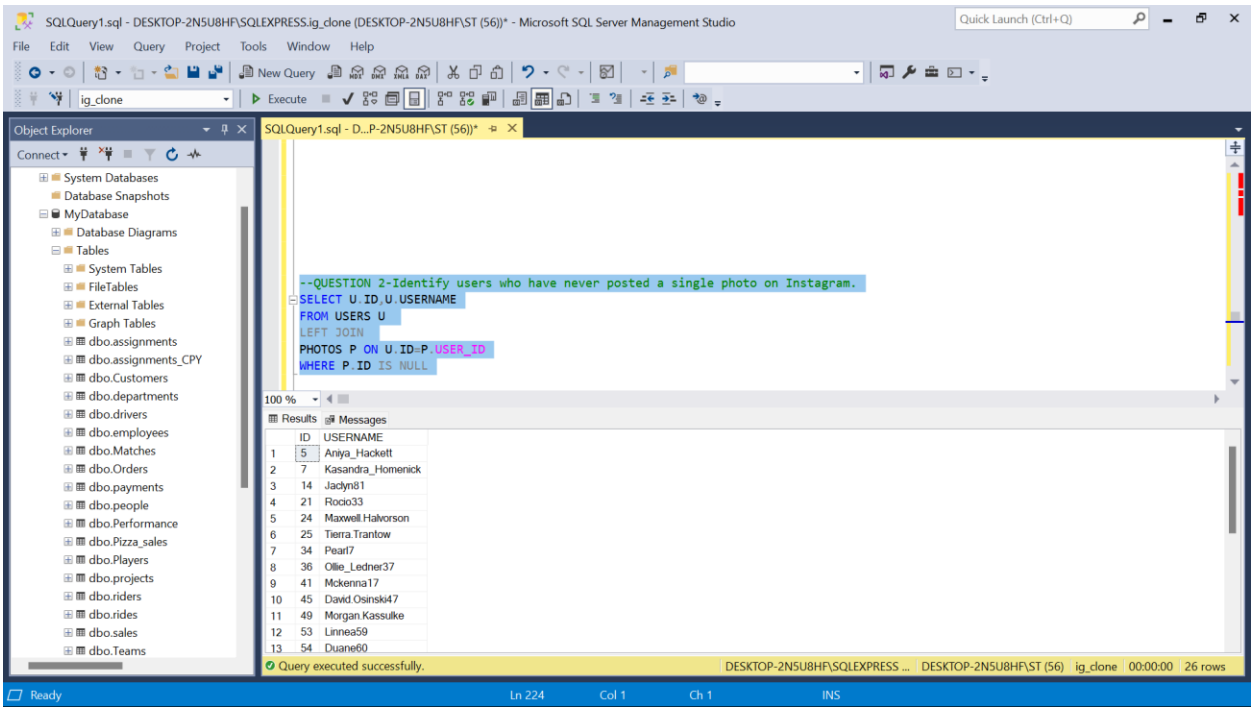
**CONCLUSION-**

**WE CAN SEND PROMOTIONAL EMAILS TO MAKE THEM ACTIVE ON THE PLATFORM ARE-**

**Aniya_Hackett,Kasandra_Homenick,Jaclyn81,Rocio33,Maxwell.Halvorson,Tierra.Trantow,Pearl7,Ollie_Ledner37,Mckenna17,**

**David.Osinski47,Morgan.Kassulke,Linnea59,Duane60,Julien_Schmidt,Mike.Auer39,Franco_Keebler64,Nia_Haag,**

**Hulda.Macejkovic,Leslie67,Janelle.Nikolaus81,Darby_Herzog,Esther.Zulauf61,Bartholome.Bernhard,Jessyca_West,Esmeralda.Mraz57,Bethany20**

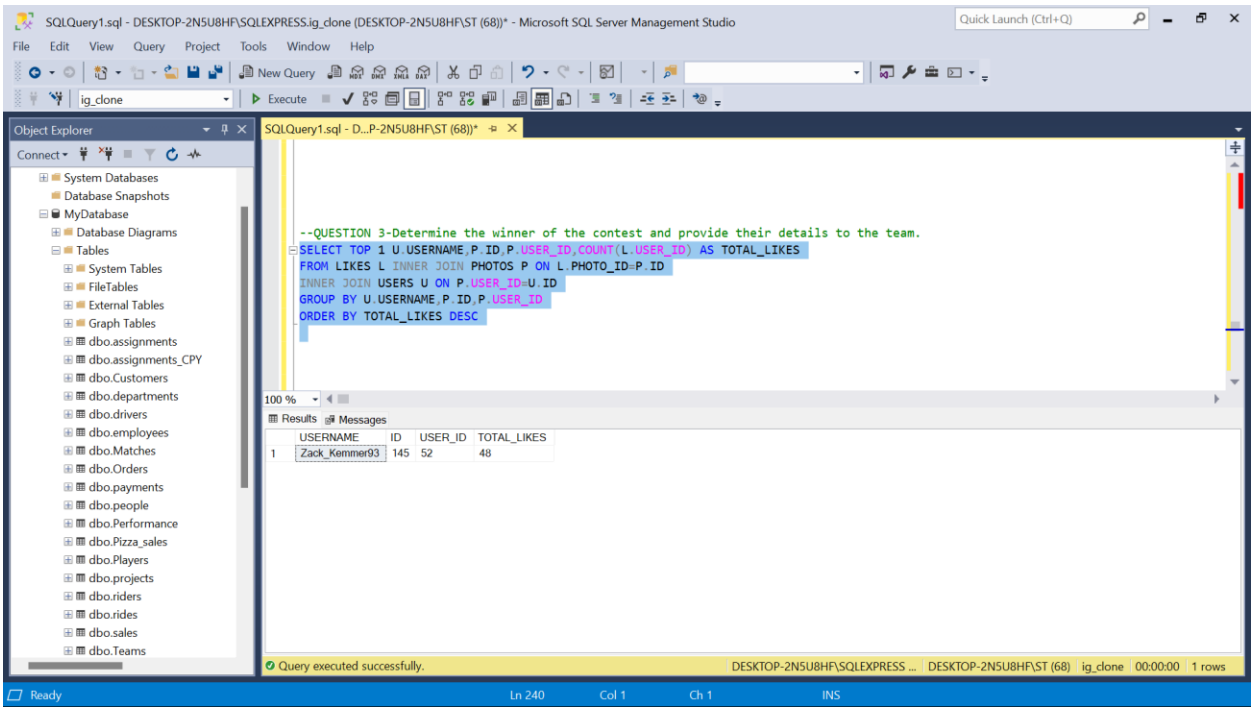



## --QUESTION 3-Determine the winner of the contest and provide their details to the team.

```
SELECT TOP 1 U.USERNAME,P.ID,P.USER_ID,COUNT(L.USER_ID) AS TOTAL_LIKES

FROM LIKES L INNER JOIN PHOTOS P ON L.PHOTO_ID=P.ID

INNER JOIN USERS U ON P.USER_ID=U.ID

GROUP BY U.USERNAME,P.ID, P.USER_ID

ORDER BY TOTAL_LIKES DESC
```

**CONCLUSION-**

**WINNER OF THE CONTEST WHO HAVE MOST LIKES ON THE SINGLE PICTURE IS - Zack_Kemmer93**

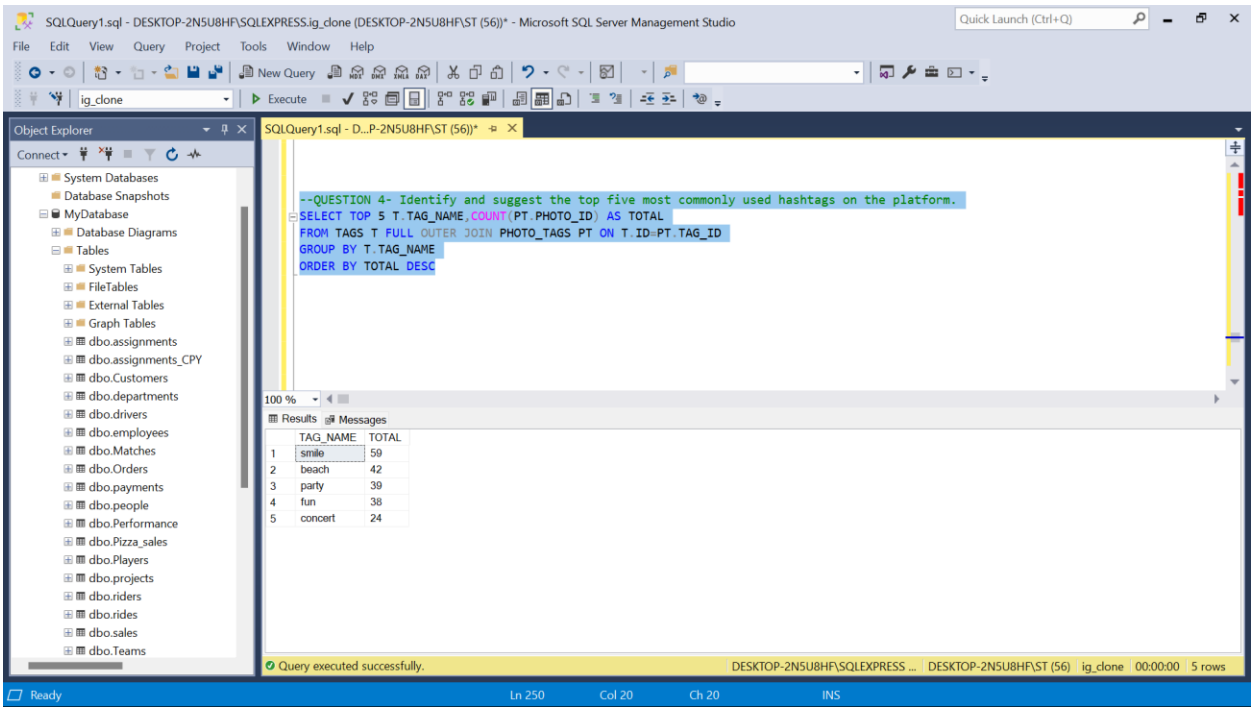



**--QUESTION 4- Identify and suggest the top five most commonly used hashtags on the platform.**

```
SELECT TOP 5 T.TAG_NAME,COUNT(PT.PHOTO_ID) AS TOTAL

FROM TAGS T FULL OUTER JOIN PHOTO_TAGS PT ON T.ID=PT.TAG_ID

GROUP BY T.TAG_NAME

ORDER BY TOTAL DESC
```

**CONCLUSION-**

**MOST POPULAR HASHTAGS USED BY THE USERS ARE- smile, beach, party, fun, concert**

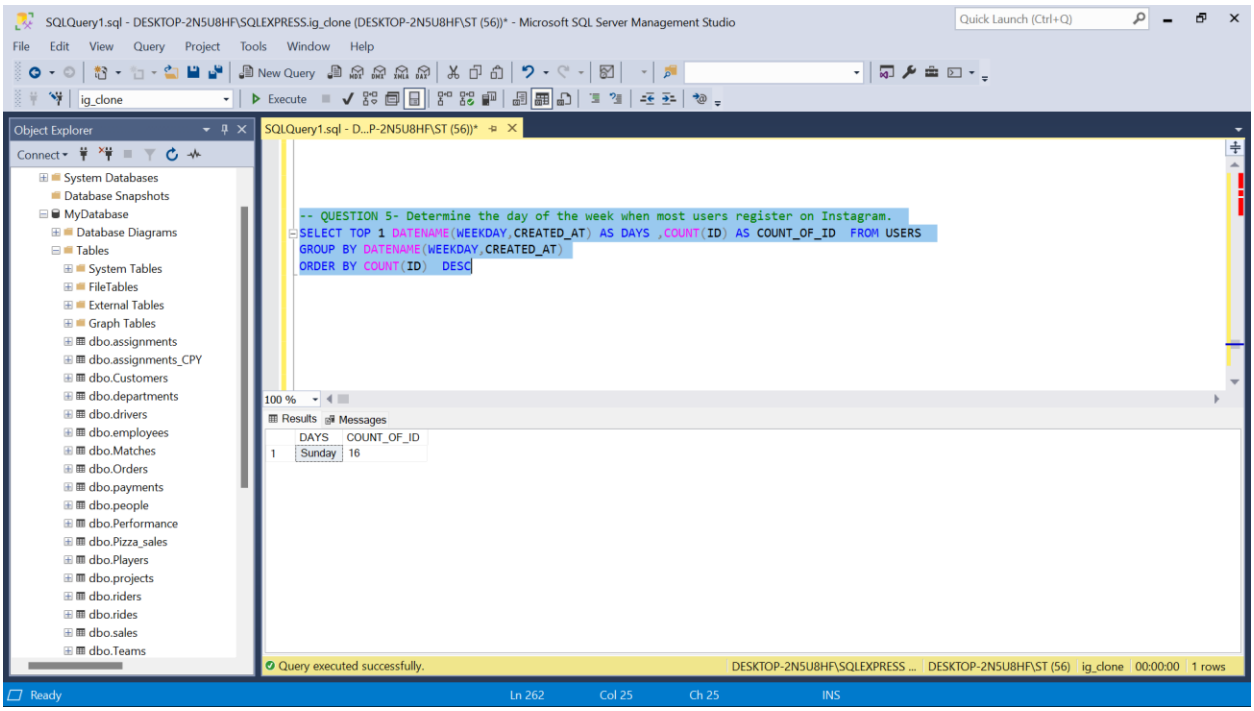



## -- QUESTION 5- Determine the day of the week when most users register on Instagram.

```
SELECT TOP 1 DATENAME(WEEKDAY,CREATED_AT) AS DAYS ,COUNT(ID) AS COUNT_OF_ID  FROM USERS

GROUP BY DATENAME(WEEKDAY,CREATED_AT)

ORDER BY COUNT(ID)  DESC
```

**CONCLUSION-**

**THE BEST DAY TO LAUNCH ADS ON THE PLATFORM IS SUNDAY AS MOST OF THE USERS ARE ACTIVE ON WEEKENDS.**




**--QUESTION 6- Calculate the average number of posts per user on Instagram. Also, provide the**

**SELECT COUNT(P.ID) AS TOTAL_POSTS,**

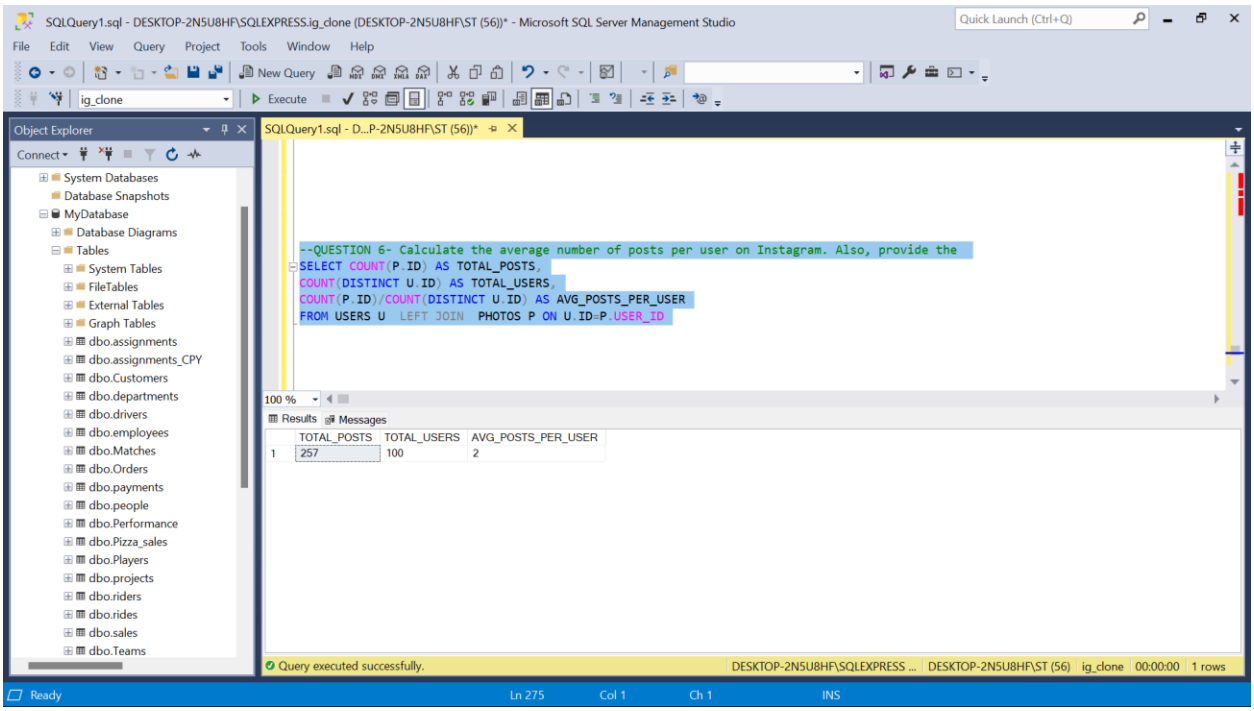
**COUNT(DISTINCT U.ID) AS TOTAL_USERS,**

**COUNT(P.ID)/COUNT(DISTINCT U.ID) AS AVG_POSTS_PER_USER**

**FROM USERS U  LEFT JOIN  PHOTOS P ON U.ID=P.USER_ID**

**CONCLUSION-**

**AVERAGE NO. OF POSTS BY EACH USER IS 2.**

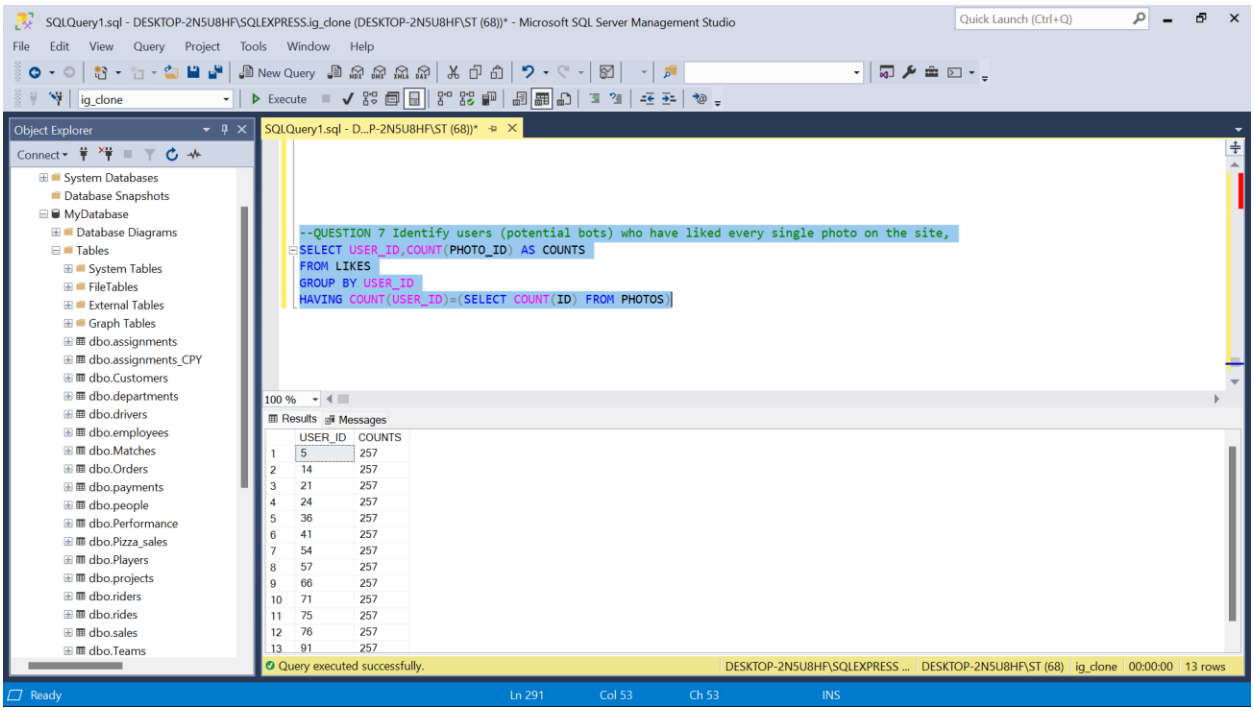



## --QUESTION 7 Identify users (potential bots) who have liked every single photo on the site.

```
SELECT USER_ID,COUNT(PHOTO_ID) AS COUNTS

FROM LIKES

GROUP BY USER_ID

HAVING COUNT(USER_ID)=(SELECT COUNT(ID) FROM PHOTOS)
```

**CONCLUSION-**

**AROUND 13 PEOPLE HAVE FAKE ACCOUNTS ON INSTAGRAM AS THEY HAVE LIKED EACH AND EVERY PHOTO ON INSTAGRAM WHICH IS NOT POSSIBLE FOR A NORMAL USERS**




## INSIGHTS-

- ✓ OLDEST USERS-Identified the platform's earliest adopters and they are the most loyal users eligible for rewards.
- ✓ INACTIVE PEOPLE-We can run some interactive Campaigns that can encourage users to become more active on the platform
- ✓ CONTEST WINNER-Determined the most engaging photo and its owner, which could inform future contests.
- ✓ HASHTAG RESEARCH-Found the most-used hash tags to optimize content creation.
- ✓ AD-CAMPAIGN LAUNCH- This information is useful for scheduling promotions and marketing efforts effectively. SUNDAY is the day when we can run some campaigns as most users spend their leisure time on social media platforms.
- ✓ USER-ENGAGEMENT-Measured user engagement by calculating average posts per user.
- ✓ FAKE ACCOUNTS DETECTION- We can block those accounts or we can give them some warnings if they are violating platform policies.

## RESULTS-

Through this project I have analyzed and provided actionable insights, such as determining the best time for ad campaigns, identifying highly active users, and detected fake accounts. These findings can help improve user engagement strategies and ensure platform integrity. Applying these insights can help the marketing team to use these insights to launch a new campaign, the product team to use them to decide on new features to build, and the development team can use them to improve the overall user experience.