



Instagram User Analytics

PROJECT DESCRIPTION:

I am Anjali Rajput, a data analyst working with the project team at Instagram. Our role involve user analysis which means tracking how user engage with the digital product. The insights derived from these analysis can be by the different teams inside the business. In this project we are using MySQL workbench as a tool to analyse and answer the question posed by the management team.

This project focus on two major factors :

1. Marketing Matrix
2. Investor Matrix



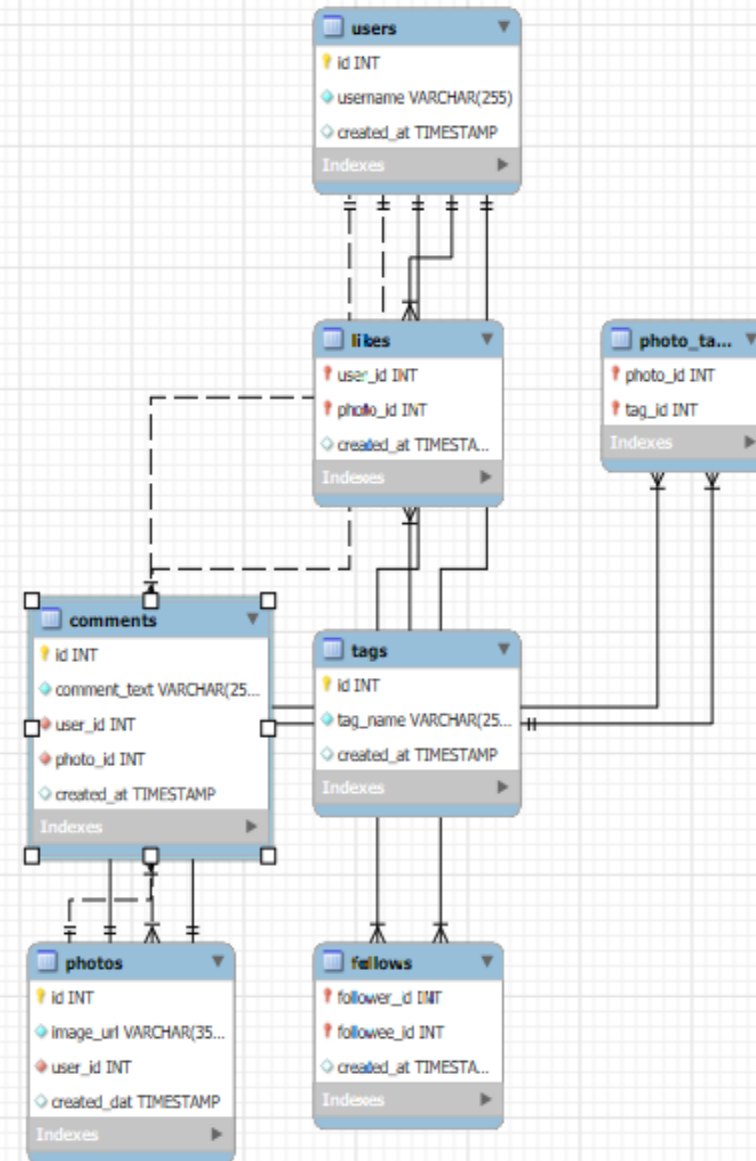
Approach:

- Database is created in MySQL workbench by executing multiple queries.
- The dataset is studied by the help of er diagram.

Tech Stack Used:

- MySQL workbench is used to create database.

ER Diagram:



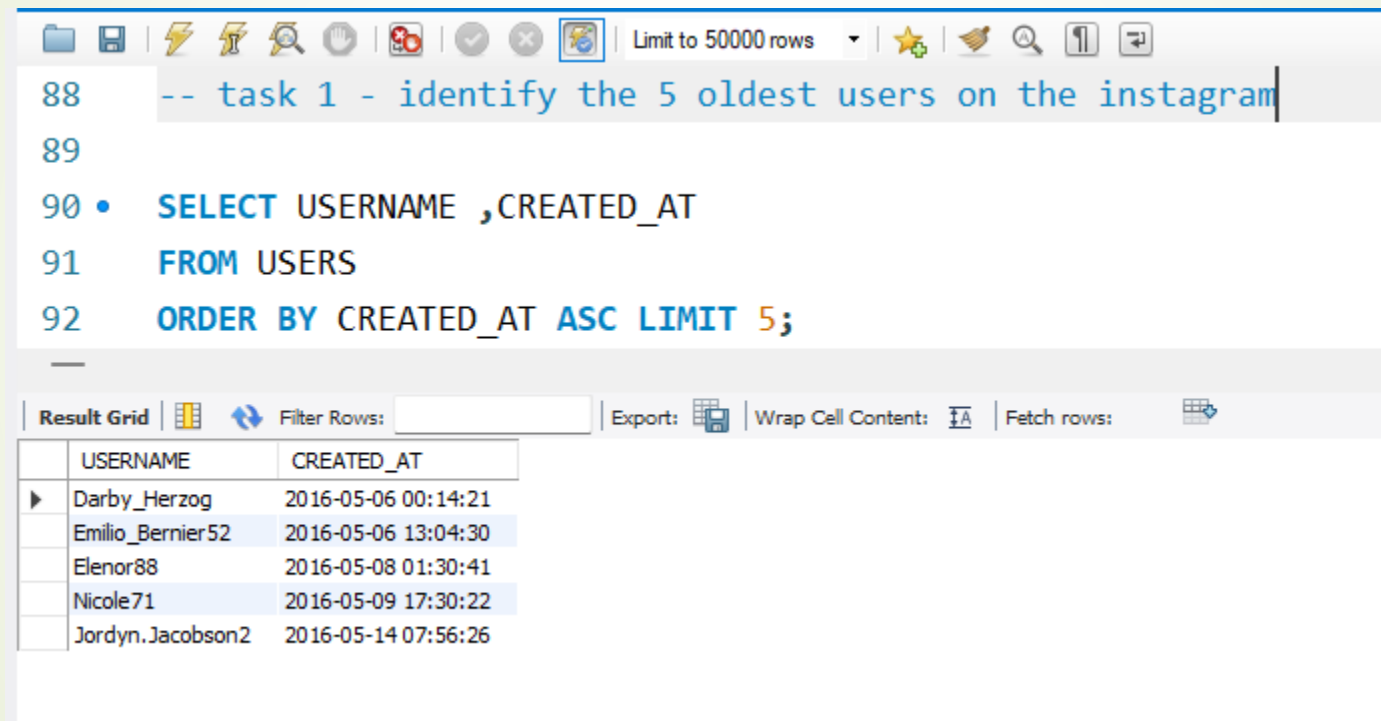


Insights



TASK 1- Loyal User Reward:

Here are top 5 most loyal users with their username and registration date.



The screenshot shows a SQL query editor interface. The query is as follows:

```
-- task 1 - identify the 5 oldest users on the instagram
SELECT USERNAME ,CREATED_AT
FROM USERS
ORDER BY CREATED_AT ASC LIMIT 5;
```

Below the query, the results are displayed in a table with columns USERNAME and CREATED_AT. The results show the 5 oldest users based on their registration date.

| 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 |

TASK 2 — Inactive User Engagement-

There are only 5 people who have not posted single photo on Instagram.

```
94      -- task 2 - identify the user who have not posted single photo on instagram
95  •   SELECT USERNAME, USERS.ID AS USERS_ID
96      FROM USERS
97      LEFT JOIN PHOTOS
98      ON USERS.ID=PHOTOS.USER_ID
99      WHERE PHOTOS.ID IS NULL
100     ORDER BY USER_ID;
101
```

Result Grid

Filter Rows:

Export:

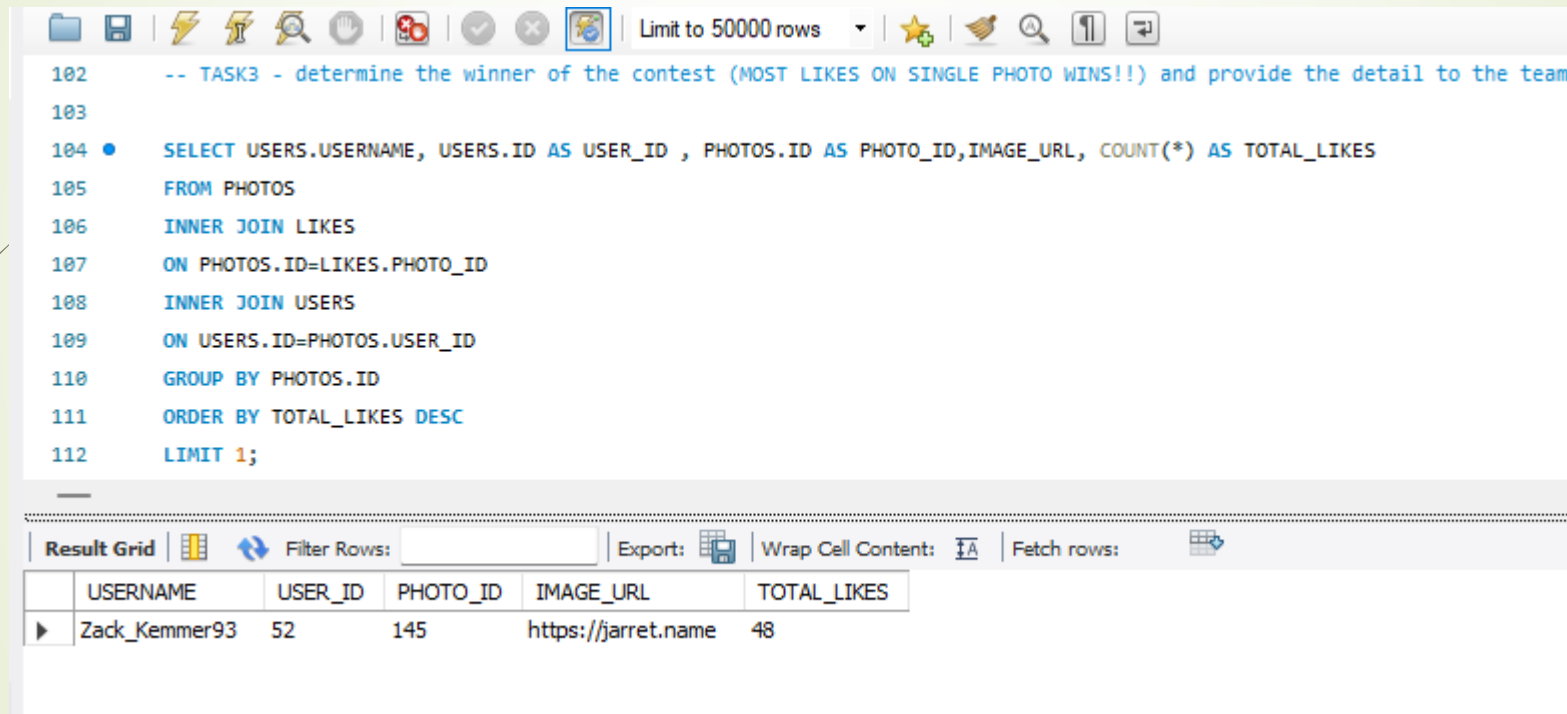
Wrap Cell Content:

Fetch rows:

| | 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 |

TASK 3- Contest Winner Declaration-

The contest Winner is Zack_Kemmer93 with user ID 52 and photo ID 145 with total of 48 likes.



The screenshot displays a SQL query editor interface. The query is as follows:

```
-- TASK3 - determine the winner of the contest (MOST LIKES ON SINGLE PHOTO WINS!!) and provide the detail to the team  
  
SELECT USERS.USERNAME, USERS.ID AS USER_ID , PHOTOS.ID AS PHOTO_ID, IMAGE_URL, COUNT(*) AS TOTAL_LIKES  
FROM PHOTOS  
INNER JOIN LIKES  
ON PHOTOS.ID=LIKES.PHOTO_ID  
INNER JOIN USERS  
ON USERS.ID=PHOTOS.USER_ID  
GROUP BY PHOTOS.ID  
ORDER BY TOTAL_LIKES DESC  
LIMIT 1;
```

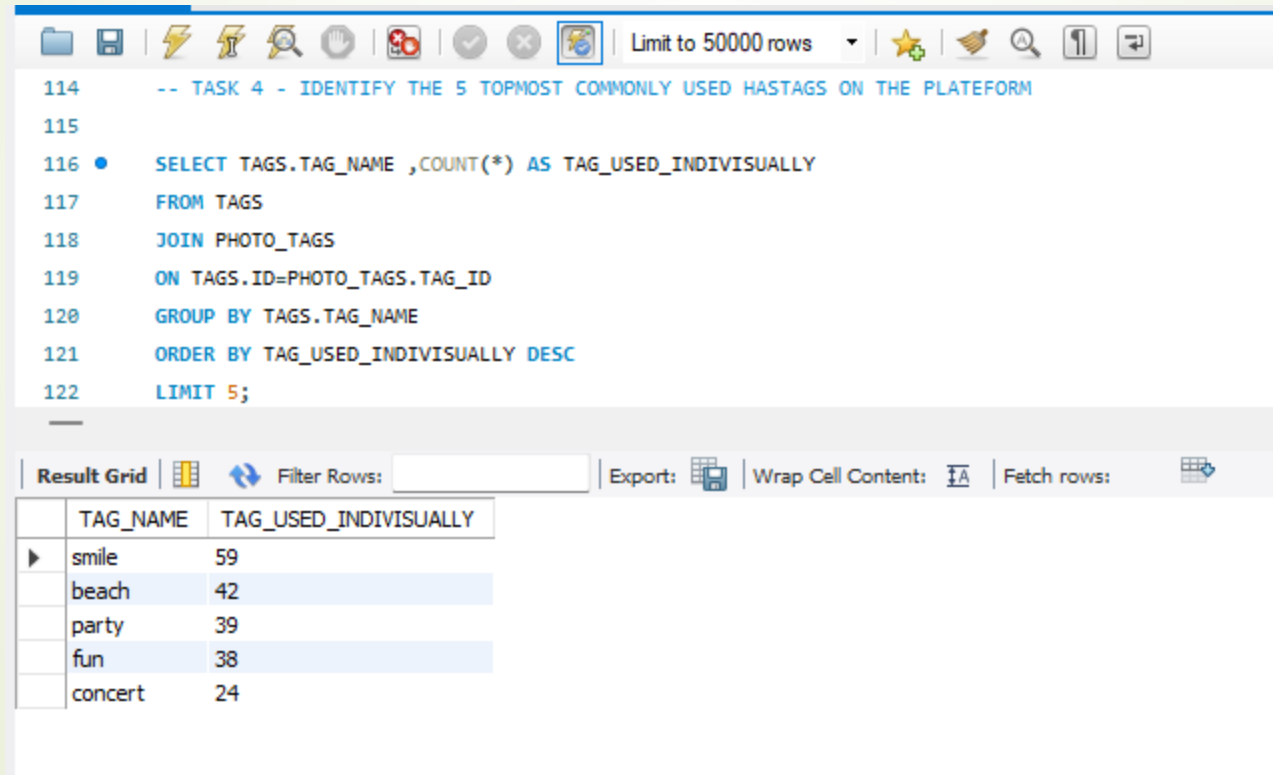
Below the query editor, the 'Result Grid' is shown, displaying the results of the query:

| USERNAME | USER_ID | PHOTO_ID | IMAGE_URL | TOTAL_LIKES |
|---------------|---------|----------|---------------------|-------------|
| Zack_Kemmer93 | 52 | 145 | https://jarret.name | 48 |

TASK 4- Hashtag Research:

Here is the list of 5 topmost commonly used Hashtags on the platform.

- Smile, -Beach, -Party, -Fun, -Concert



The screenshot displays a SQL query editor interface. The query is as follows:

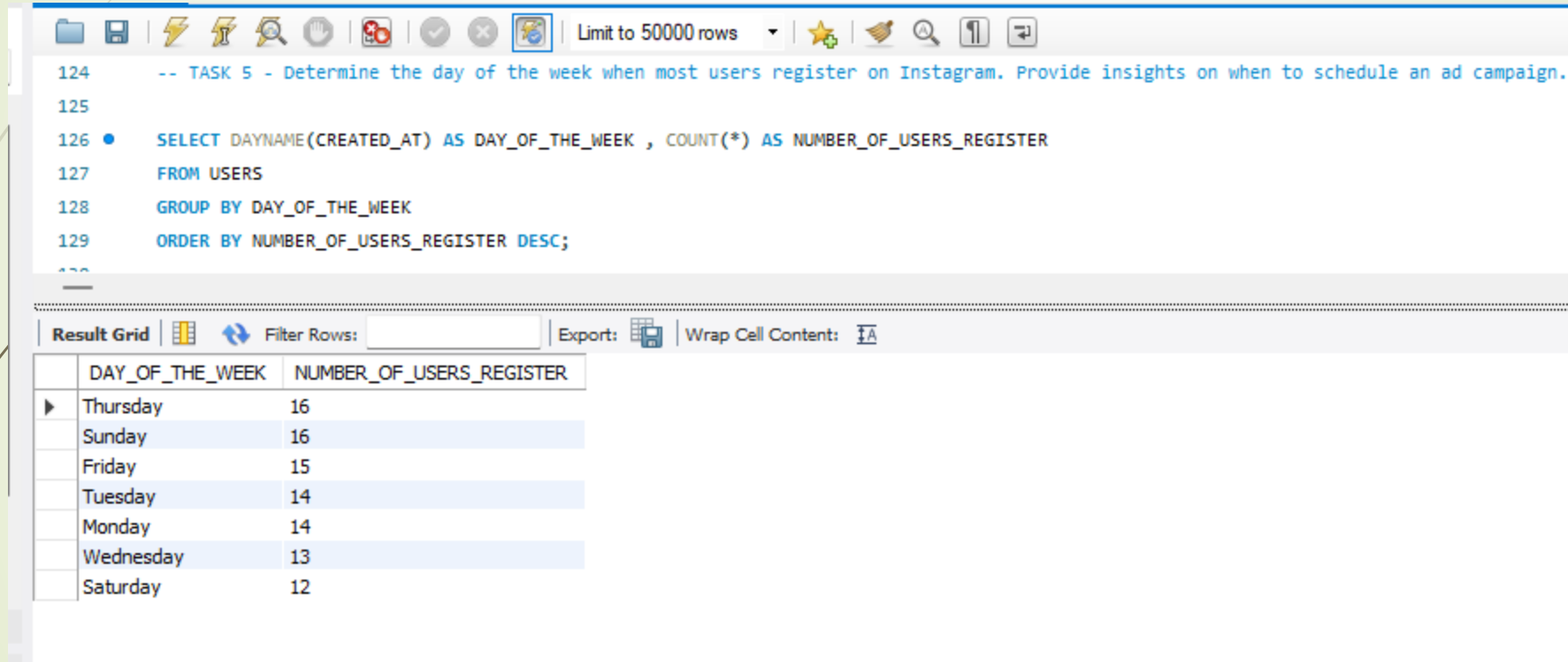
```
114 -- TASK 4 - IDENTIFY THE 5 TOPMOST COMMONLY USED HASTAGS ON THE PLATFORM
115
116 • SELECT TAGS.TAG_NAME ,COUNT(*) AS TAG_USED_INDIVISUALLY
117 FROM TAGS
118 JOIN PHOTO_TAGS
119 ON TAGS.ID=PHOTO_TAGS.TAG_ID
120 GROUP BY TAGS.TAG_NAME
121 ORDER BY TAG_USED_INDIVISUALLY DESC
122 LIMIT 5;
```

Below the query editor, the 'Result Grid' shows the following data:

| | TAG_NAME | TAG_USED_INDIVISUALLY |
|---|----------|-----------------------|
| ▶ | smile | 59 |
| | beach | 42 |
| | party | 39 |
| | fun | 38 |
| | concert | 24 |

TASK 5- Add Campaign Launch-

Thursday and Saturday are the days of the week when most user registered.



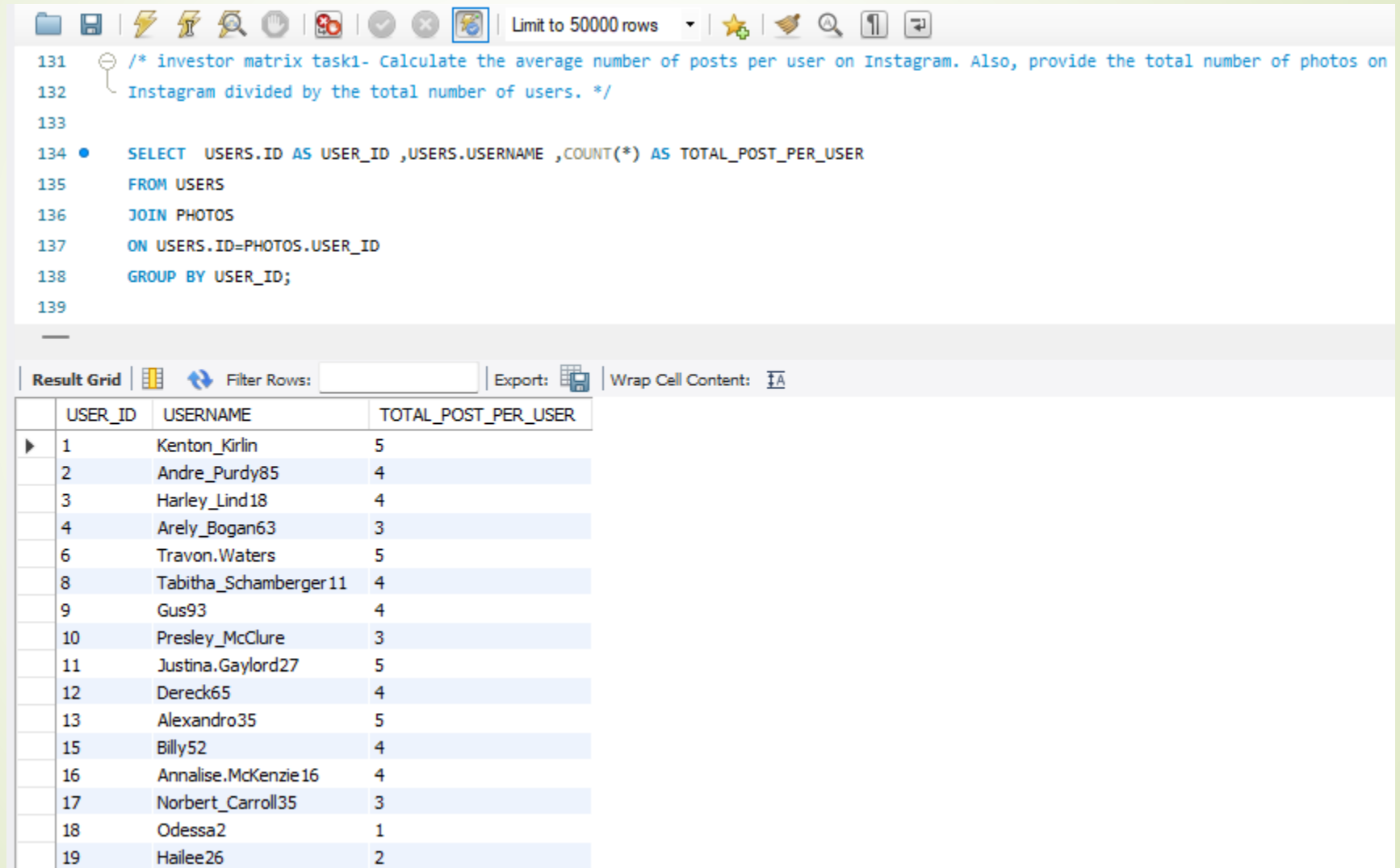
```
124 -- TASK 5 - Determine the day of the week when most users register on Instagram. Provide insights on when to schedule an ad campaign.
125
126 • SELECT DAYNAME(CREATED_AT) AS DAY_OF_THE_WEEK , COUNT(*) AS NUMBER_OF_USERS_REGISTER
127 FROM USERS
128 GROUP BY DAY_OF_THE_WEEK
129 ORDER BY NUMBER_OF_USERS_REGISTER DESC;
```

Result Grid

| | DAY_OF_THE_WEEK | NUMBER_OF_USERS_REGISTER |
|---|-----------------|--------------------------|
| ▶ | Thursday | 16 |
| | Sunday | 16 |
| | Friday | 15 |
| | Tuesday | 14 |
| | Monday | 14 |
| | Wednesday | 13 |
| | Saturday | 12 |

TASK 1- User Engagement:

the average number of post per user on Instagram.



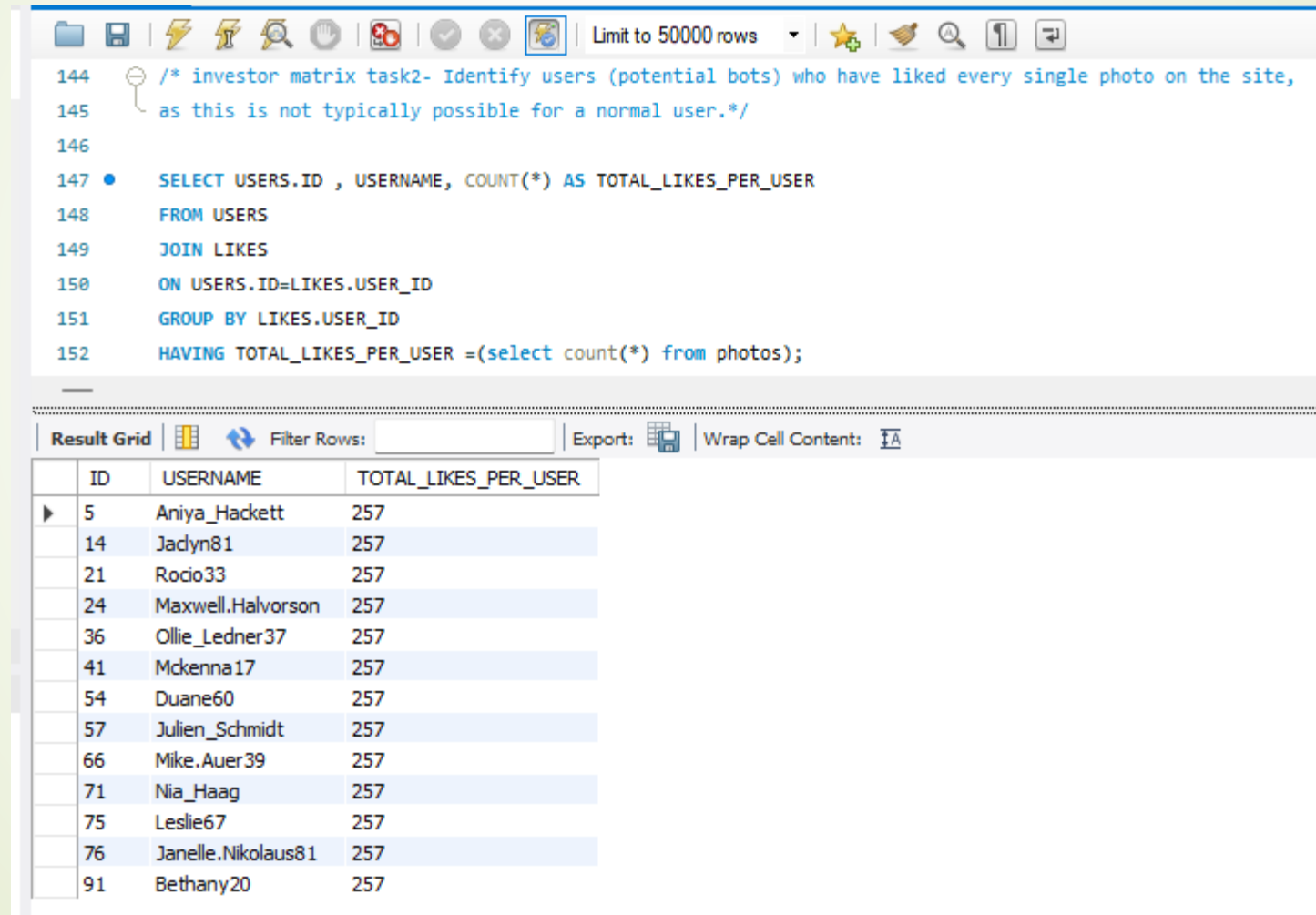
```
131  /* investor matrix task1- Calculate the average number of posts per user on Instagram. Also, provide the total number of photos on
132  Instagram divided by the total number of users. */
133
134  ●  SELECT  USERS.ID AS USER_ID ,USERS.USERNAME ,COUNT(*) AS TOTAL_POST_PER_USER
135      FROM  USERS
136      JOIN  PHOTOS
137      ON    USERS.ID=PHOTOS.USER_ID
138      GROUP BY USER_ID;
139
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: [IA](#)

| | USER_ID | USERNAME | TOTAL_POST_PER_USER |
|---|---------|-----------------------|---------------------|
| ▶ | 1 | Kenton_Kirlin | 5 |
| | 2 | Andre_Purdy85 | 4 |
| | 3 | Harley_Lind18 | 4 |
| | 4 | Arely_Bogan63 | 3 |
| | 6 | Travon.Waters | 5 |
| | 8 | Tabitha_Schamberger11 | 4 |
| | 9 | Gus93 | 4 |
| | 10 | Presley_McClure | 3 |
| | 11 | Justina.Gaylord27 | 5 |
| | 12 | Dereck65 | 4 |
| | 13 | Alexandro35 | 5 |
| | 15 | Billy52 | 4 |
| | 16 | Annalise.McKenzie16 | 4 |
| | 17 | Norbert_Carroll35 | 3 |
| | 18 | Odessa2 | 1 |
| | 19 | Hailee26 | 2 |

TASK 2- Bots & Fake Accounts:

There are 13 people who have liked each and every post on Instagram.



```
144 /* investor matrix task2- Identify users (potential bots) who have liked every single photo on the site,
145 as this is not typically possible for a normal user.*/
146
147 • SELECT USERS.ID , USERNAME, COUNT(*) AS TOTAL_LIKES_PER_USER
148 FROM USERS
149 JOIN LIKES
150 ON USERS.ID=LIKES.USER_ID
151 GROUP BY LIKES.USER_ID
152 HAVING TOTAL_LIKES_PER_USER =(select count(*) from photos);
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: [IA](#)

| | ID | USERNAME | TOTAL_LIKES_PER_USER |
|---|----|--------------------|----------------------|
| ▶ | 5 | Aniya_Hackett | 257 |
| | 14 | Jadyn81 | 257 |
| | 21 | Rocio33 | 257 |
| | 24 | Maxwell.Halvorson | 257 |
| | 36 | Ollie_Ledner37 | 257 |
| | 41 | Mckenna17 | 257 |
| | 54 | Duane60 | 257 |
| | 57 | Julien_Schmidt | 257 |
| | 66 | Mike.Auer39 | 257 |
| | 71 | Nia_Haag | 257 |
| | 75 | Leslie67 | 257 |
| | 76 | Janelle.Nikolaus81 | 257 |
| | 91 | Bethany20 | 257 |



RESULT:

- **SQL Proficiency** : Develop a strong command on SQL queries .
- **Data exploration** : Navigate complex databases to extract relevant information.
- **Data Mining** : Use E-R Diagram for mining data.
- **Business Decision Making** : I learned about how few insights can have a huge impact on business decision making process.

Also learned how real life business case scenarios are.



THANK YOU !