Instagram User Analytics

Description:

User analysis is the process by which we track how users engage and interact with our digital product (software or mobile application) in an attempt to derive business insights for marketing, product & development teams.
These insights are then used by teams across the business to launch a new marketing campaign, decide on features to build for an app, track the success of the app by measuring user engagement and improve the experience altogether while helping the business grow.
You are working with the product team of Instagram and the product manager has asked you to provide insights on the questions asked by the management team.

Tech stack used:

This project is executed using MySQL workbench.

Approach and Insights:

We have performed analysis using SQL to answer the questions asked by the management team .

A) Marketing: The marketing team wants to launch some campaigns, and they need your help with the following

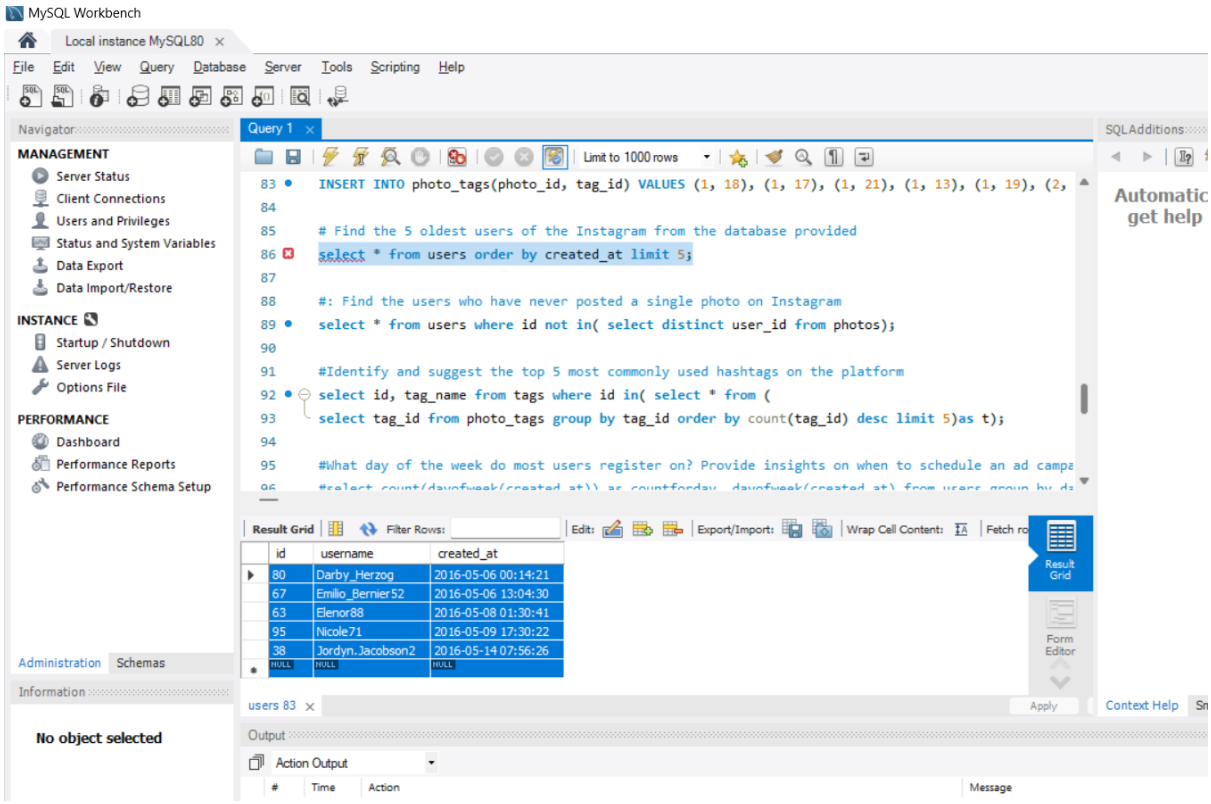
1. **Rewarding Most Loyal Users:** People who have been using the platform for the longest time.
   Your Task: Find the 5 oldest users of the Instagram from the database provided

Query :

```
select * from users order by created_at limit 5;
```

Output :

80	Darby_Herzog	2016-05-06 00:14:21
67	Emilio_Bernier52	2016-05-06 13:04:30
63	Elenor88	2016-05-08 01:30:41
95	Nicole71	2016-05-09 17:30:22
38	Jordyn.Jacobson2	2016-05-14 07:56:26



The 5 oldest users of instagram are :

- Darby_Herzog
- Emilio_Bernier52
- Elenor88
- Nicole71
- Jordyn.Jacobson2

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

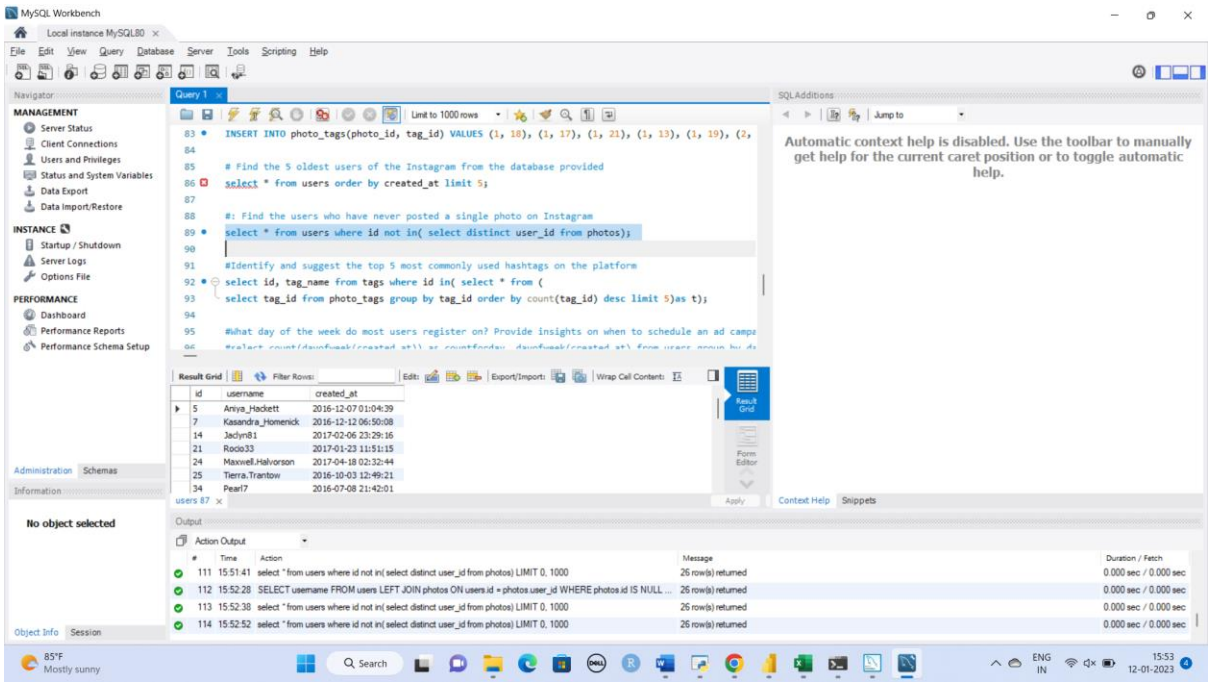
Query:

```
select * from users where id not in( select distinct user_id from photos);
```

Output:

5	Aniya_Hackett	2016-12-07 01:04:39
7	Kasandra_Homenick	2016-12-12 06:50:08
14	Jaclyn81	2017-02-06 23:29:16

21 Rocio33	2017-01-23 11:51:15
24 Maxwell.Halvorson	2017-04-18 02:32:44
25 Tierra.Trantow	2016-10-03 12:49:21
34 Pearl7	2016-07-08 21:42:01
36 Ollie_Ledner37	2016-08-04 15:42:20
41 Mckenna17	2016-07-17 17:25:45
45 David.Osinski47	2017-02-05 21:23:37
49 Morgan.Kassulke	2016-10-30 12:42:31
53 Linnea59	2017-02-07 07:49:34
54 Duane60	2016-12-21 04:43:38
57 Julien_Schmidt	2017-02-02 23:12:48
66 Mike.Auer39	2016-07-01 17:36:15
68 Franco_Keebler64	2016-11-13 20:09:27
71 Nia_Haag	2016-05-14 15:38:50
74 Hulda.Macejkovic	2017-01-25 17:17:28
75 Leslie67	2016-09-21 05:14:01
76 Janelle.Nikolaus81	2016-07-21 09:26:09
80 Darby_Herzog	2016-05-06 00:14:21
81 Esther.Zulauf61	2017-01-14 17:02:34
83 Bartholome.Bernhard	2016-11-06 02:31:23
89 Jessyca_West	2016-09-14 23:47:05
90 Esmeralda.Mraz57	2017-03-03 11:52:27
91 Bethany20	2016-06-03 23:31:53



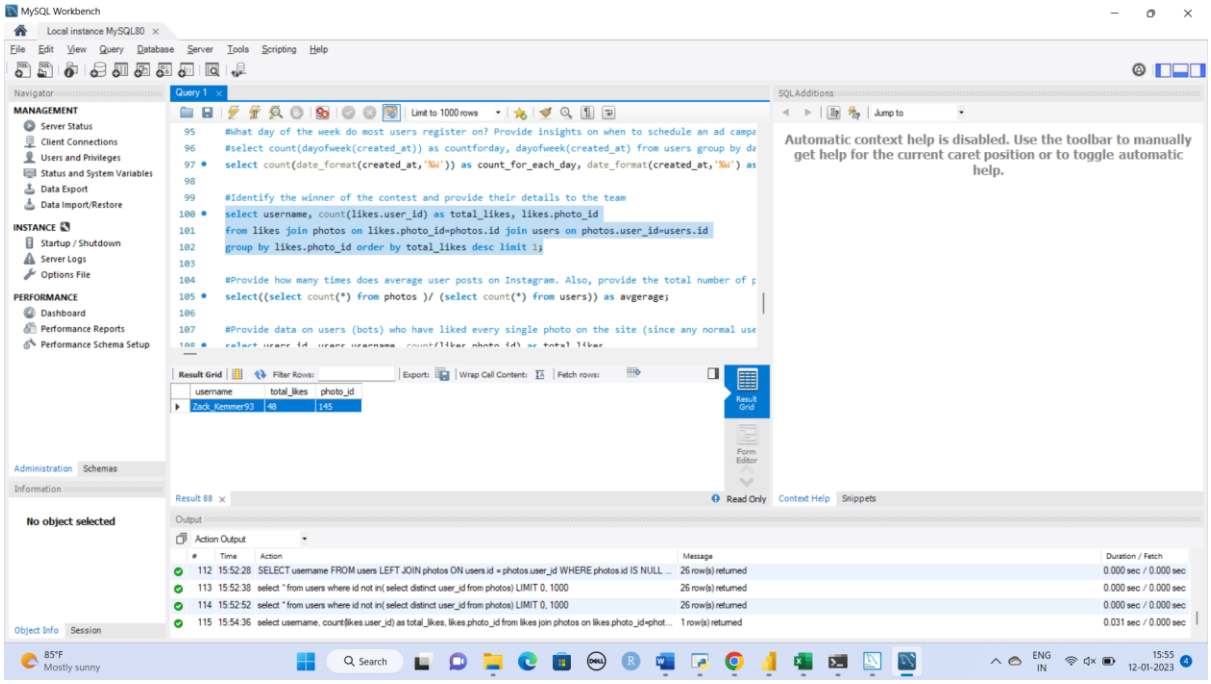
3. **Declaring Contest Winner:** The team started a contest and the user who gets the most likes on a single photo will win the contest now they wish to declare the winner.
   Your Task: Identify the winner of the contest and provide their details to the team.

Query:

```
select username, count(likes.user_id) as total_likes, likes.photo_id
from likes join photos on likes.photo_id=photos.id join users on photos.user_id=users.id
group by likes.photo_id order by total_likes desc limit 1;
```

Output:

Zack_Kemmer93 48 145



Zack_Kemmer93 is the winner of the contest with 48 likes for his photo with id.

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

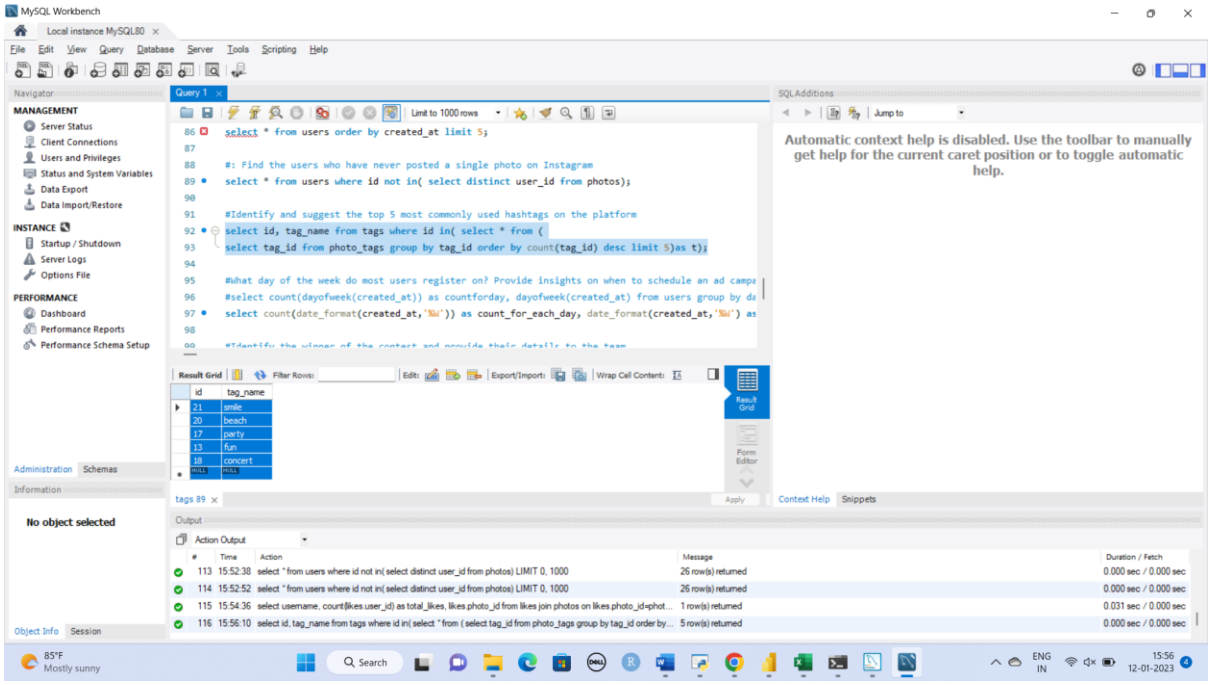
Query:

```
select id, tag_name from tags where id in( select * from (

select tag_id from photo_tags group by tag_id order by count(tag_id) desc limit 5)as t);
```

Output :

21 smile
20 beach
17 party
13 fun
18 concert



The top 5 most commonly used hashtags in the platform are :

- smile
- beach
- party
- fun
- concert

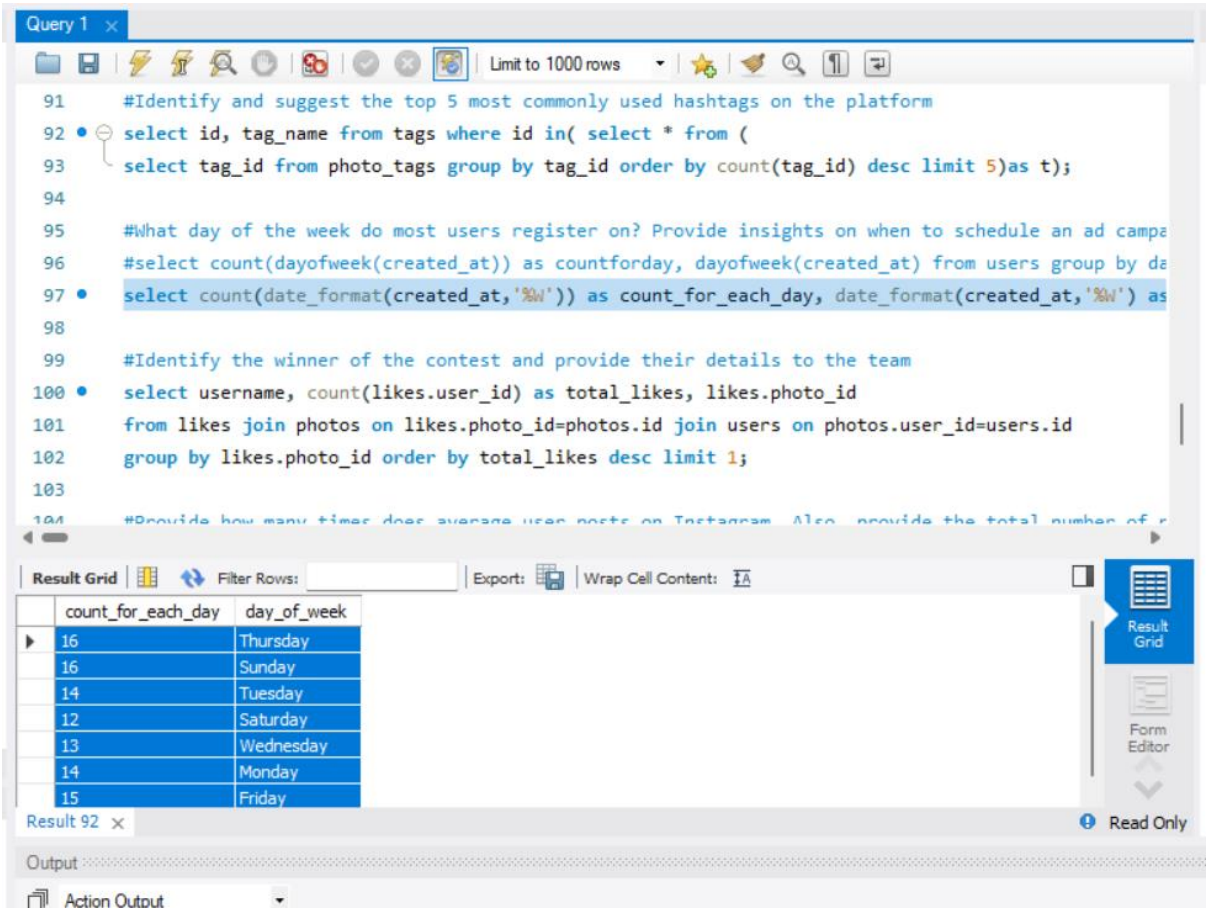
5. **Launch AD Campaign:** The team wants to know, which day would be the best day to launch ADs.
   Your Task: What day of the week do most users register on? Provide insights on when to schedule an ad campaign

Query :

```
select count(date_format(created_at,'%W')) as count_for_each_day,
date_format(created_at,'%W') as day_of_week from users group by dayofweek(created_at);
```

Output:

16 Thursday
16 Sunday
14 Tuesday
12 Saturday
13 Wednesday
14 Monday
15 Friday



From the analysis we find that most users register on Thursday and Sunday. So it is most likely the best to launch the Ads on Thursday and Sunday.

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

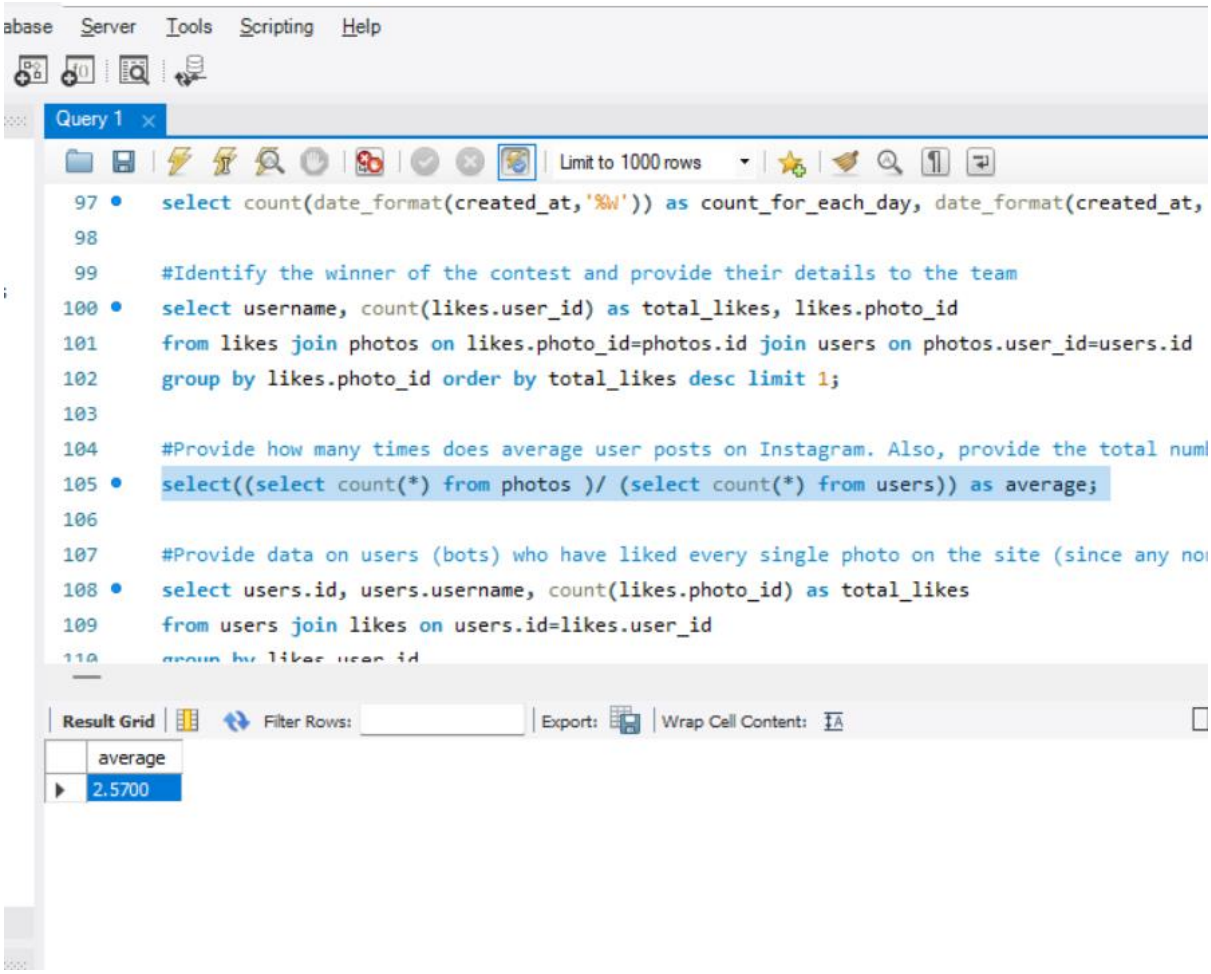
Your Task: Provide how many times does average user posts on Instagram. Also, provide the total number of photos on Instagram/total number of users

Query :

```
select((select count(*) from photos )/ (select count(*) from users)) as average;
```

Output :

2.5700



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).

Query :

```
select users.id, users.username, count(likes.photo_id) as total_likes

from users join likes on users.id=likes.user_id
```

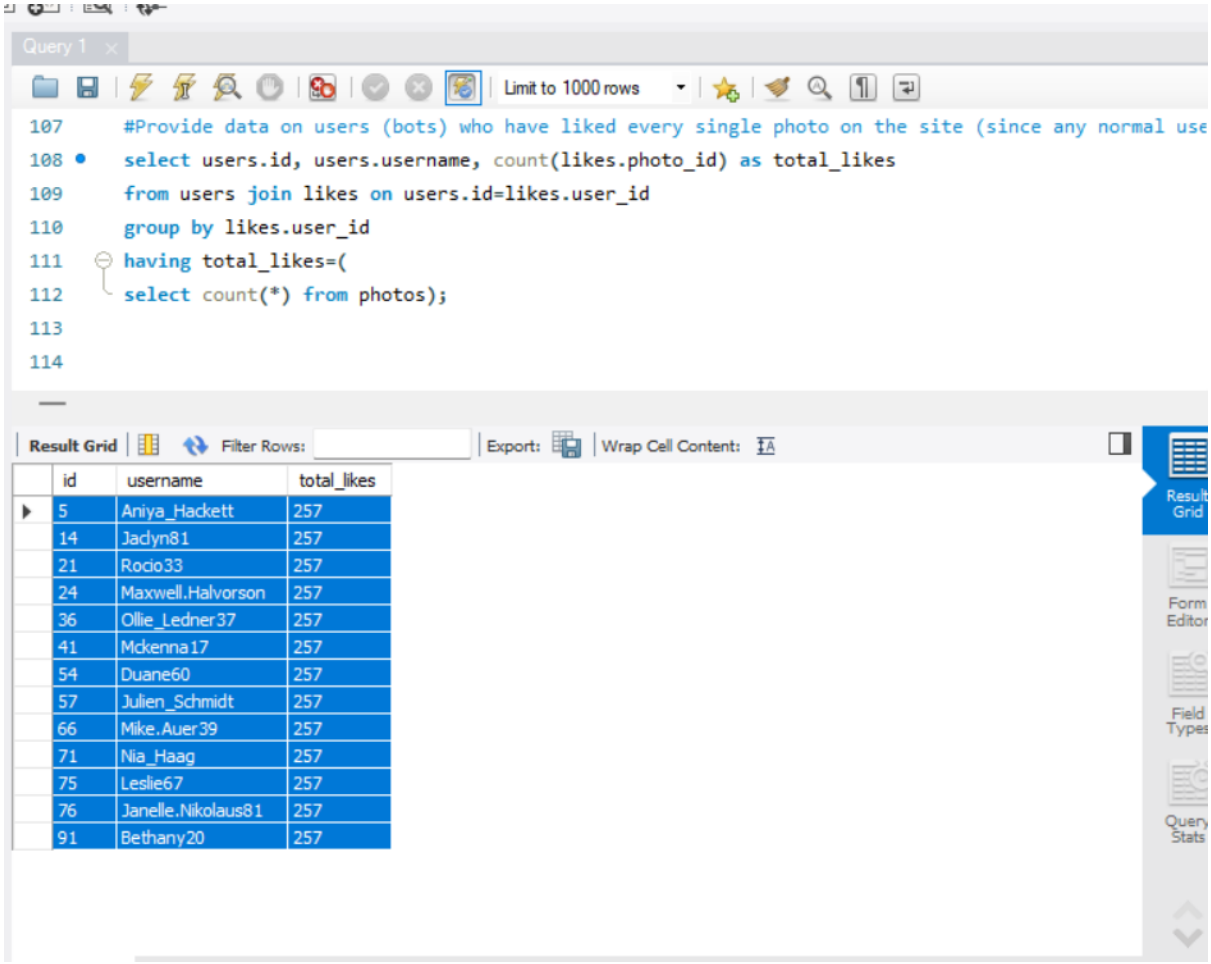
```
group by likes.user_id

having total_likes=(

select count(*) from photos);
```

Output :

5	Aniya_Hackett	257
14	Jaclyn81	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 :

By working on this case study I have learnt a lot about SQL and how to apply the SQL queries to answer questions and get insights from the analysis.