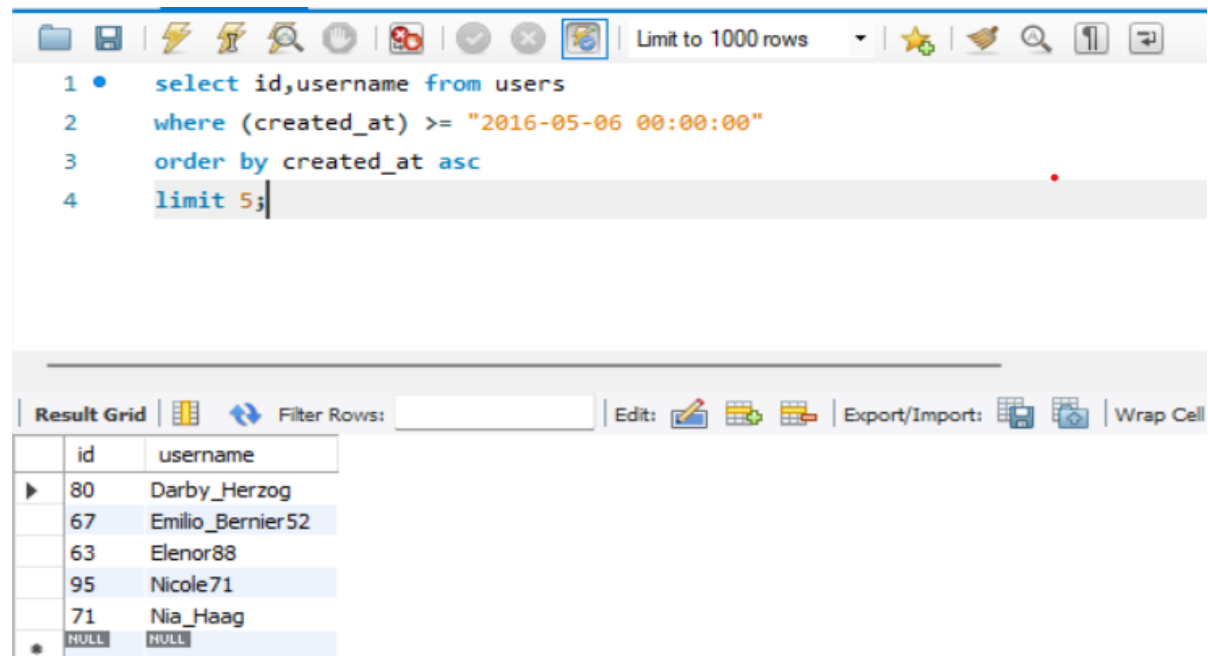


Explanation: As a analyst, I have to do the analysis the data of the Instagram users to get the grow in their business.

Approach: Firstly, I write all the questions that I want a solution of. Then I perform SQL queries.

Firstly, to analyze whom users using their account for long time. In short I have to find the top 5 oldest users. For this I write the below sql query:

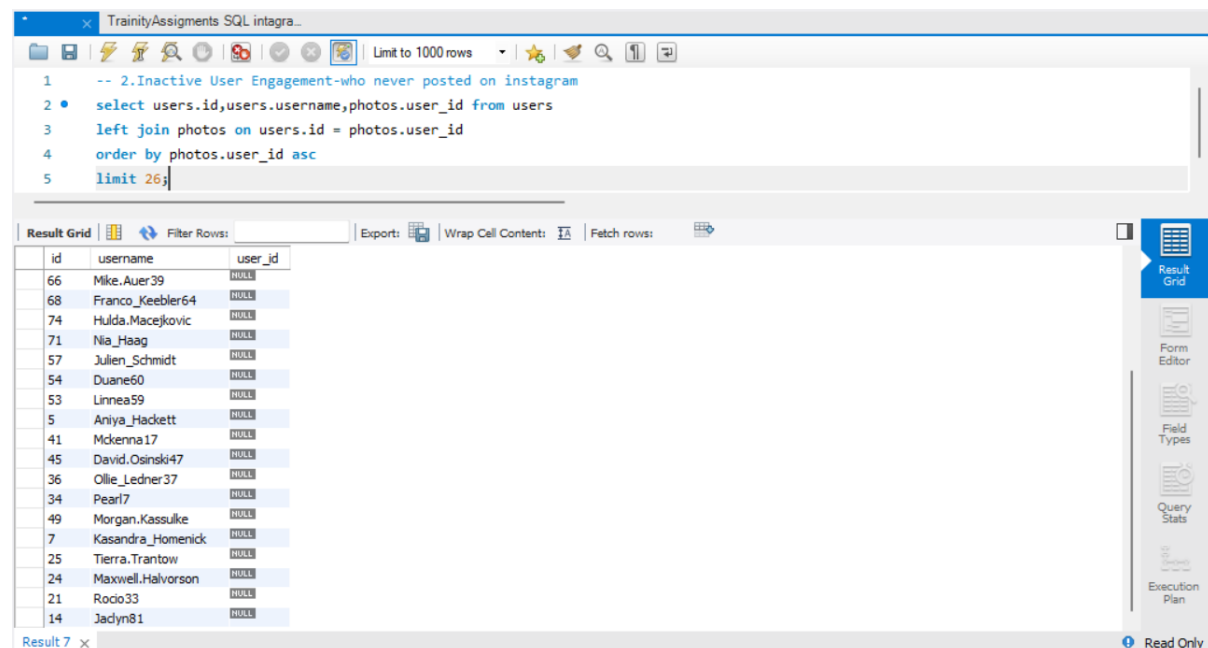


```
1 • select id,username from users
2   where (created_at) >= "2016-05-06 00:00:00"
3   order by created_at asc
4   limit 5;
```

Result Grid

	id	username
▶	80	Darby_Herzog
	67	Emilio_Bernier52
	63	Elenor88
	95	Nicole71
	71	Nia_Haag
*	NULL	NULL

The second question is I have to find the users who never posted on Instagram and the sql query is:



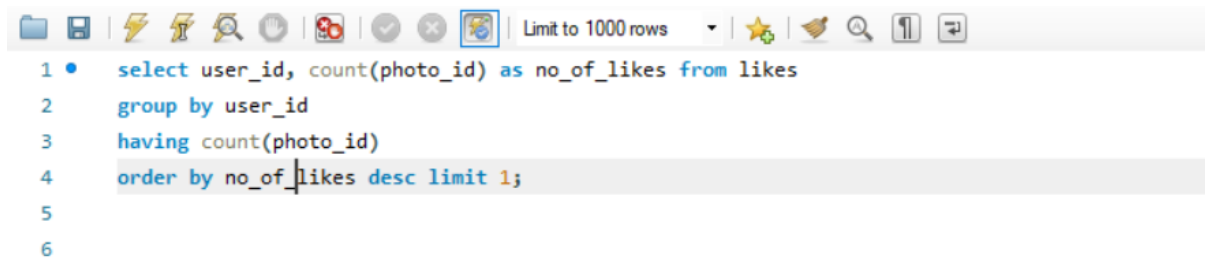
```
1 -- 2.Inactive User Engagement-who never posted on instagram
2 • select users.id,users.username,photos.user_id from users
3   left join photos on users.id = photos.user_id
4   order by photos.user_id asc
5   limit 26;
```

Result Grid

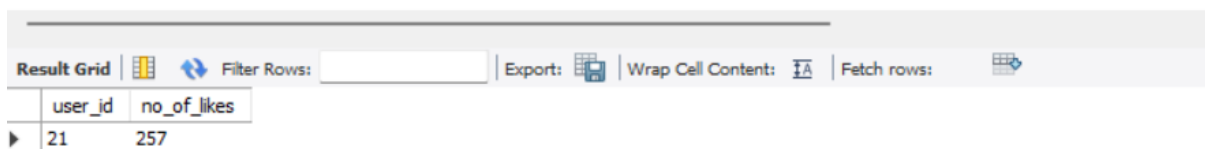
	id	username	user_id
	66	Mike_Auer39	NULL
	68	Franco_Keebler64	NULL
	74	Hulda.Macejkovic	NULL
	71	Nia_Haag	NULL
	57	Julien_Schmidt	NULL
	54	Duane60	NULL
	53	Linnea59	NULL
	5	Aniya_Hackett	NULL
	41	Mckenna17	NULL
	45	David.Osinski47	NULL
	36	Ollie_Ledner37	NULL
	34	Pearl7	NULL
	49	Morgan.Kassulke	NULL
	7	Kassandra_Homenick	NULL
	25	Tierra.Trantow	NULL
	24	Maxivell.Halvorson	NULL
	21	Rocio33	NULL
	14	Jadyn81	NULL

The username who have null values never posted on Instagram.

The third question is Contest Winner Declaration-most likes on single photo and sql query for this question is:



```
1 • select user_id, count(photo_id) as no_of_likes from likes
2   group by user_id
3   having count(photo_id)
4   order by no_of_likes desc limit 1;
5
6
```



	user_id	no_of_likes
▶	21	257

In the above screenshot we can see that user\_id – 21 have the most likes on their photo.

I used group by to separate the data user-wise and I used order-by because I have to see the no of likes and “desc” means the no of likes are shown in descending order means we can see the most of likes first.

Fourth question is Identify and suggest the top five most commonly used hashtags on the platform.

So the sql query for this question is:

TrainityAssignments SQL intagra...

Limit to 1000 rows

```

1 • select tag_id, count(*) as tags from photo_tags
2 group by tag_id
3 order by tags desc
4 limit 5;
5
6

```

Result Grid

Filter Rows:

Export:

Wrap Cell Content:

Fetch rows:

tag_id	tags
21	59
20	42
17	39
13	38
18	24

Here we can see that the host used hashtags are shown above with tag\_id.

I used group by to separate the data tag wise and I used order by to put tags in descending order and used limit 5 because it shows me the top 5 tags used by most of the users.

The fifth question is determine the day of the week when most users register on Instagram.

And the sql query is:

TrainityAssignments SQL intagra...

Limit to 1000 rows

```

1 -- 5. Ad Campaign Launch: The team wants to know the best day of the week to launch ads.
2 -- Your Task: Determine the day of the week when most users register on Instagram.
3 • SELECT created_At , COUNT(*) AS num_registrations
4 FROM users
5 GROUP BY created_at
6 ORDER BY num_registrations desc
7 limit 9;

```

Result Grid

Filter Rows:

Export:

Wrap Cell Content:

Fetch rows:

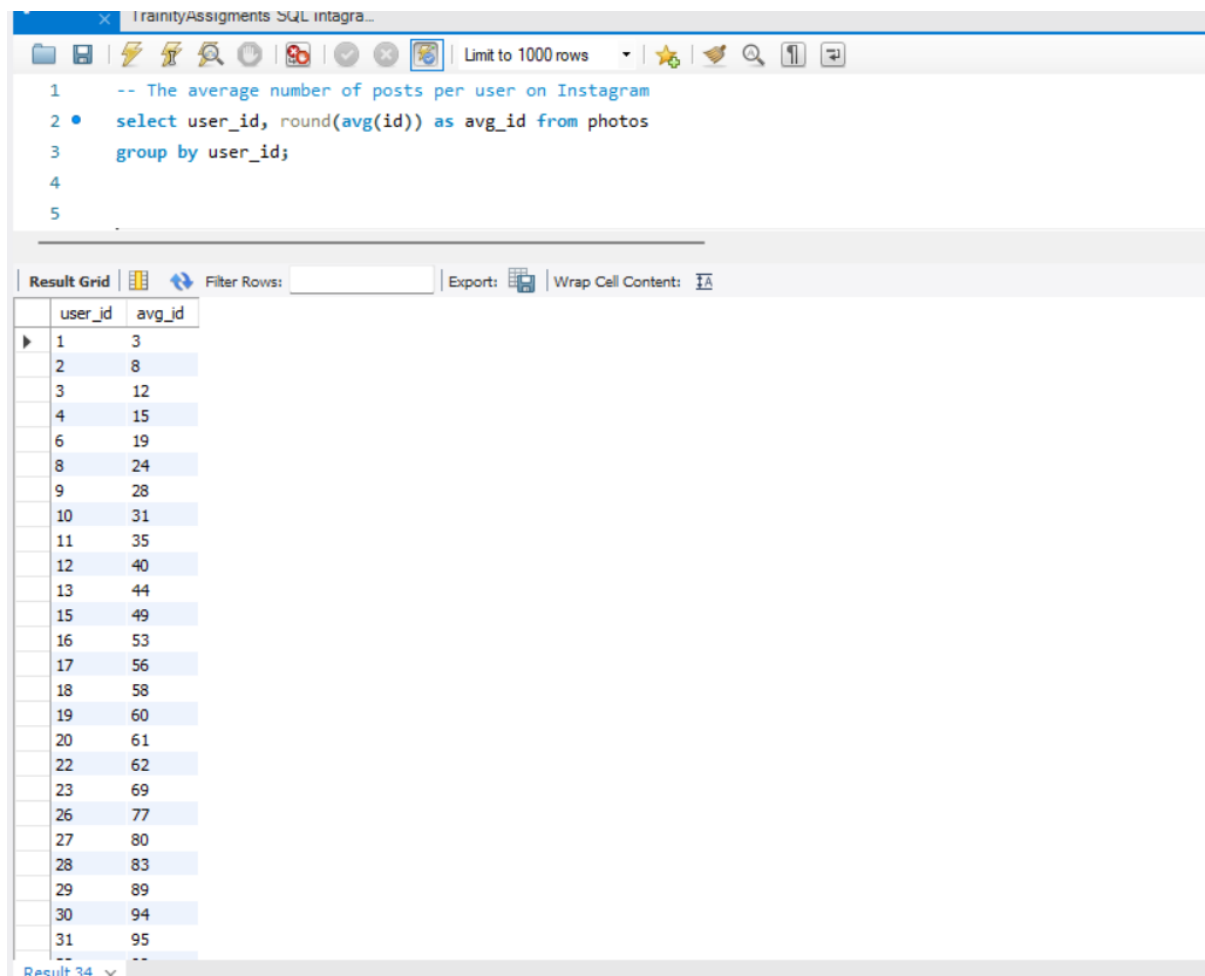
created_At	num_registrations
2016-05-14 00:00:00	2
2017-01-01 00:00:00	2
2017-03-30 00:00:00	2
2016-05-06 00:00:00	2
2017-04-30 00:00:00	2
2016-10-06 00:00:00	2
2017-01-23 00:00:00	2
2016-06-24 00:00:00	2
2017-02-06 00:00:00	2

In the SQL query we can see that these are the top 9 date when most of the users created their account.

I used group by to group the data by dates and order by desc to show the maximum no. of registrations by the date and I limit it by 9 to show only first 9 records.

The sixth question is to 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.

The first image we can see the average no of posts per user on Instagram.



1 -- The average number of posts per user on Instagram  
2 • select user\_id, round(avg(id)) as avg\_id from photos  
3 group by user\_id;  
4  
5

user_id	avg_id
1	3
2	8
3	12
4	15
6	19
8	24
9	28
10	31
11	35
12	40
13	44
15	49
16	53
17	56
18	58
19	60
20	61
22	62
23	69
26	77
27	80
28	83
29	89
30	94
31	95

In the below image we can see the total no of photos on Instagram divided by the no of users.

The screenshot shows a MySQL Workbench interface. The SQL editor contains the following query:

```
6
7
8  -- The total number of photos on Instagram divided by the total number of users.
9  • SELECT total_photos / total_users AS photos_per_user_ratio
10 FROM (
11     SELECT COUNT(*) AS total_photos
12     FROM photos
13 ) AS photo_count,
14 (
15     SELECT COUNT(*) AS total_users
16     FROM users
17 ) AS user_count;
18
19
```

Below the editor, the 'Result Grid' tab is active, showing a single column 'photos\_per\_user\_ratio' with a value of 2.5700.

The seventh question is to find the user who liked every photo on Instagram and those users can be considered as fake users.

The screenshot shows a MySQL Workbench interface. The SQL editor contains the following query:

```
1  -- 7. Bots and fake accounts - identify users (potential bots) who have liked every single photo on the site, as this is not typically p
2  • select distinct(users.id),users.username from likes
3  left join users
4  on likes.user_id = users.id;
```

Below the editor, the 'Result Grid' tab is active, showing a table with two columns: 'id' and 'username'. The results are as follows:

	id	username
2	Andre_Purdy85	
5	Aniya_Hackett	
9	Gus93	
10	Presley_McClure	
11	Justina_Gaylord27	
14	Jadyn81	
19	Hailee26	
21	Rocio33	
24	Maxwell_Halvorson	
35	Lennie_Hartmann40	
36	Ollie_Ledner37	
41	Mckenna17	

I used left join because I want to find the users who liked every photo.

3.Tech stack: I used sql language and MYSQL workbench 8.0 to run my sql queries. I used workbench 8.0 because it's the latest version. I used mysql because I used to write sql queries in workbench as a beginner.

4. Insights: By answering all the questions we get our most loyal users and at the ad campaign

They are going to get rewards on that day when most of the users registered. We can get new ideas from those users who are inactive and the owners want to know that what the users want them to be an active user.

By evaluating the user with most likes on single photo can get a reward and it will encourage those inactive users to be an active user.

5.Result: By this analysis, the rate of active users increased along with the posts per user increased by 10 %. And by branding through Instagram by the active or users with more followers they reached to their target audiences which profits the Instagram as well as user and the brand.