Instagram User Analytics Project Report

Submitted By

Soumyadip Ghosh Date-10/08/2024

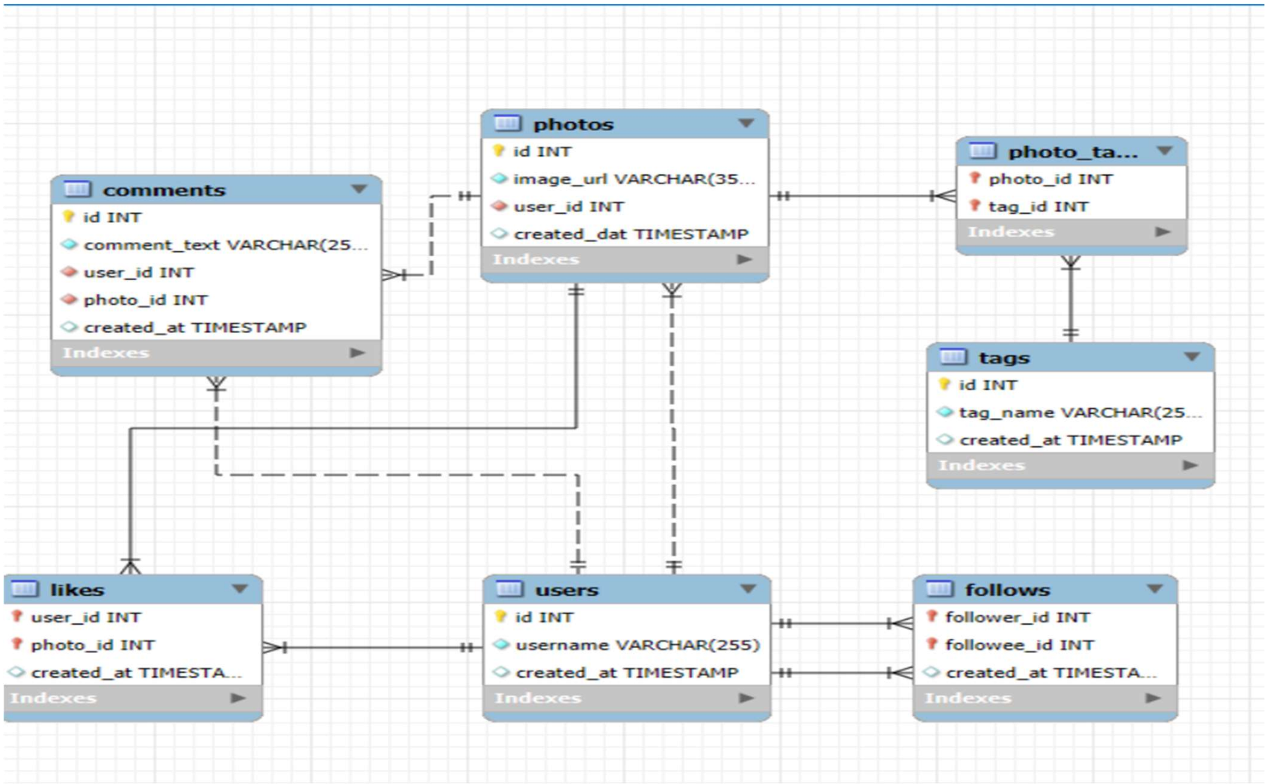
<u>1.Project Description</u> → So, basically, this project is based on analyzing Instagram users' activity. In this project, I will analyze user activity records/data and, based on the data, I will try to find some insights that will help the organization to make decisions to improve the company's growth.

<u>2.Approach</u> →To implement this project, I followed some steps. First, I checked whether the data is processable or not. Once I confirmed that the data was actionable, I started to study it. After studying the dataset, I checked the the relationships between the different tables in the database. Then, I started thinking about what insights I could retrieve from the data that would help the Marketing Team or Stakeholders to make decisions.

<u>3.Tech Stack Used</u> → In this project, I used MySQL Workbench for managing the relational database and SQL as the scripting language. I have used MySQL Workbench because I am proficient with this Database. And it quite easy and user friendly rather than other Database IDE. Ans it's provide lots off features that makes my work easy. And SQL is very easy and human being understandable scripting language that's why I used This.

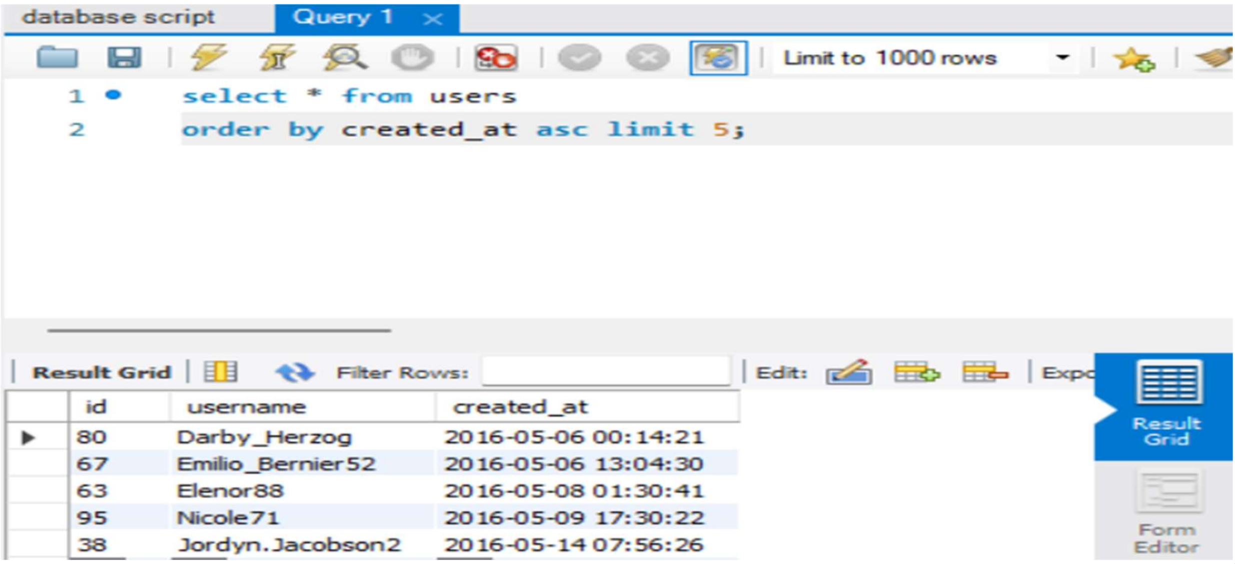
<u>4.Insights</u> → So, After all study and analyse using my Database Analysis Knowledge I have find out some Insights. So in this Project I have got some Insights where I have did Marketing Analyse and ALSO I tried to find out useful insights which will be beneficial as per Investor's Metrix.

So, The These are the Insight's and schema model of Database.



A. ***Loyal User Reward:*** *The marketing team wants to reward the most loyal users, i.e., those who have been using the platform for the longest time.*

So I find out 5 oldest user's of Instagram as per the DATASET.



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

id	username	created_at
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

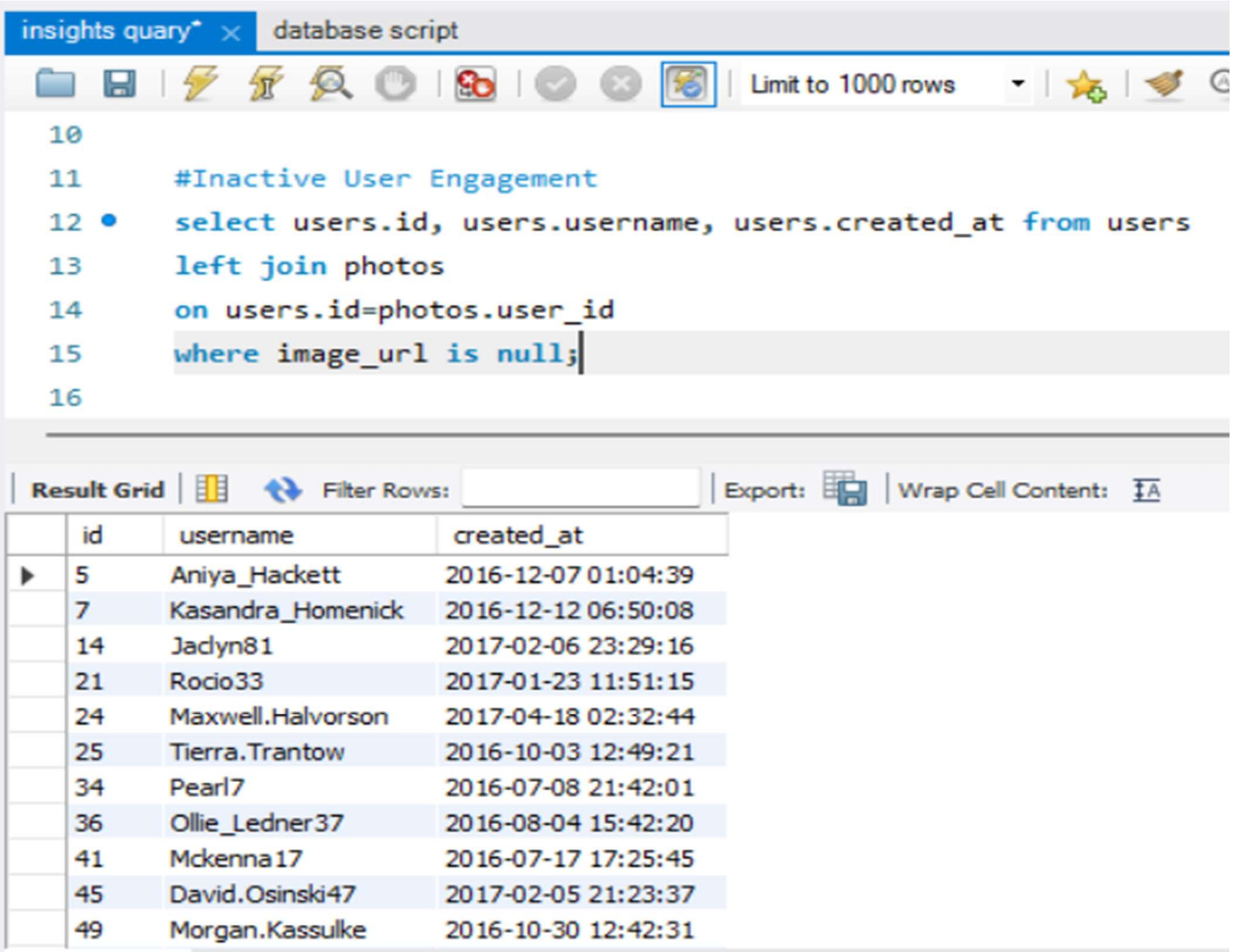
;

select * from users

order by created_at asc limit 5;

B. Inactive User Engagement: *The team wants to encourage inactive users to start posting by sending them promotional emails.*

So I find the user's who didn't posted a single photo.



```
#Inactive User Engagement
select users.id, users.username, users.created_at from users
left join photos
on users.id=photos.user_id
where image_url is null;
```

id	username	created_at
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

**** There is some data which is not in this iamge.****

Select users.id, users.username, users.created_at from users

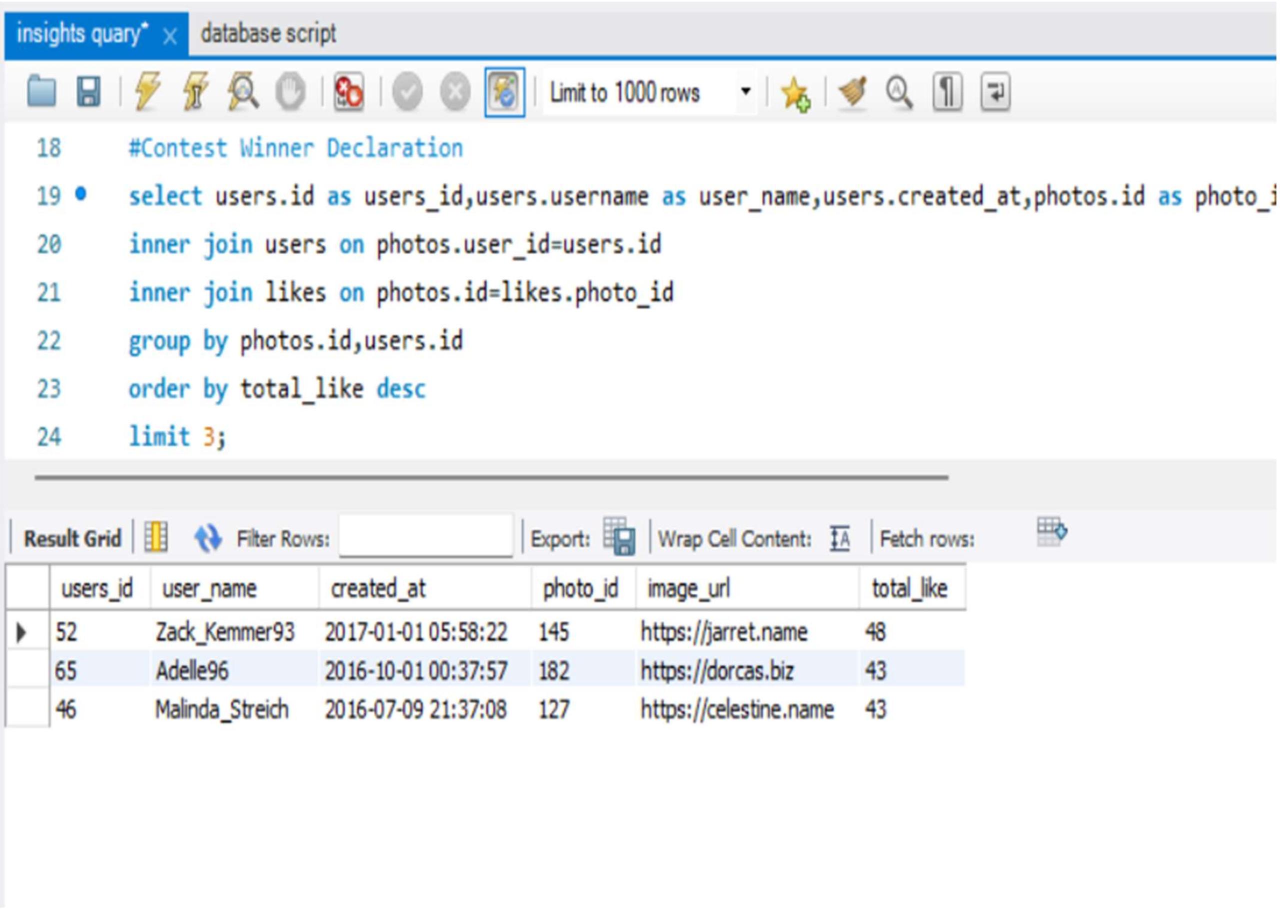
left join photos

on users.id=photos.user_id

where image_url is null;

C. ***Contest Winner Declaration:*** *The team has organized a contest where the user with the most likes on a single photo wins.*

So, I find out the 3 photo's and the user details of the photo , in which photo audience are liked most.



```
#Contest Winner Declaration
select users.id as users_id,users.username as user_name,users.created_at,photos.id as photo_i
inner join users on photos.user_id=users.id
inner join likes on photos.id=likes.photo_id
group by photos.id,users.id
order by total_like desc
limit 3;
```

users_id	user_name	created_at	photo_id	image_url	total_like
52	Zack_Kemmer93	2017-01-01 05:58:22	145	https://jarret.name	48
65	Adelle96	2016-10-01 00:37:57	182	https://dorcas.biz	43
46	Malinda_Streich	2016-07-09 21:37:08	127	https://celestine.name	43

select users.id as users_id,users.username as user_name,users.created_at,photos.id as photo_id,photos.image_url,count(likes.photo_id) as total_like from photos

inner join users on photos.user_id=users.id

inner join likes on photos.id=likes.photo_id

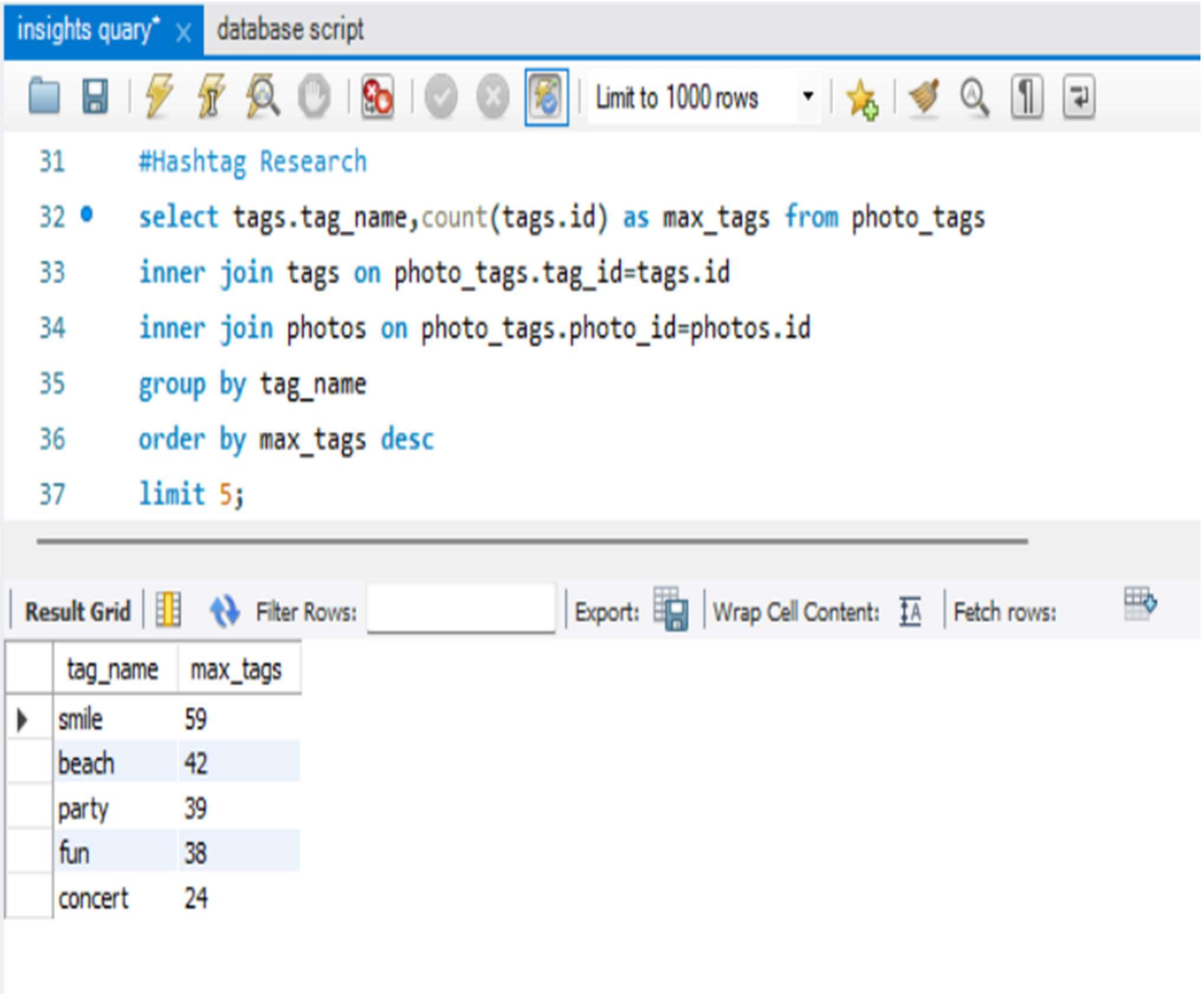
group by photos.id,users.id

order by total_like desc

limit 3;

D. Hashtag Research: *A partner brand wants to know the most popular hashtags to use in their posts to reach the most people.*

Now I find the 5 most popular Tags.



```
#Hashtag Research
select tags.tag_name,count(tags.id) as max_tags from photo_tags
inner join tags on photo_tags.tag_id=tags.id
inner join photos on photo_tags.photo_id=photos.id
group by tag_name
order by max_tags desc
limit 5;
```

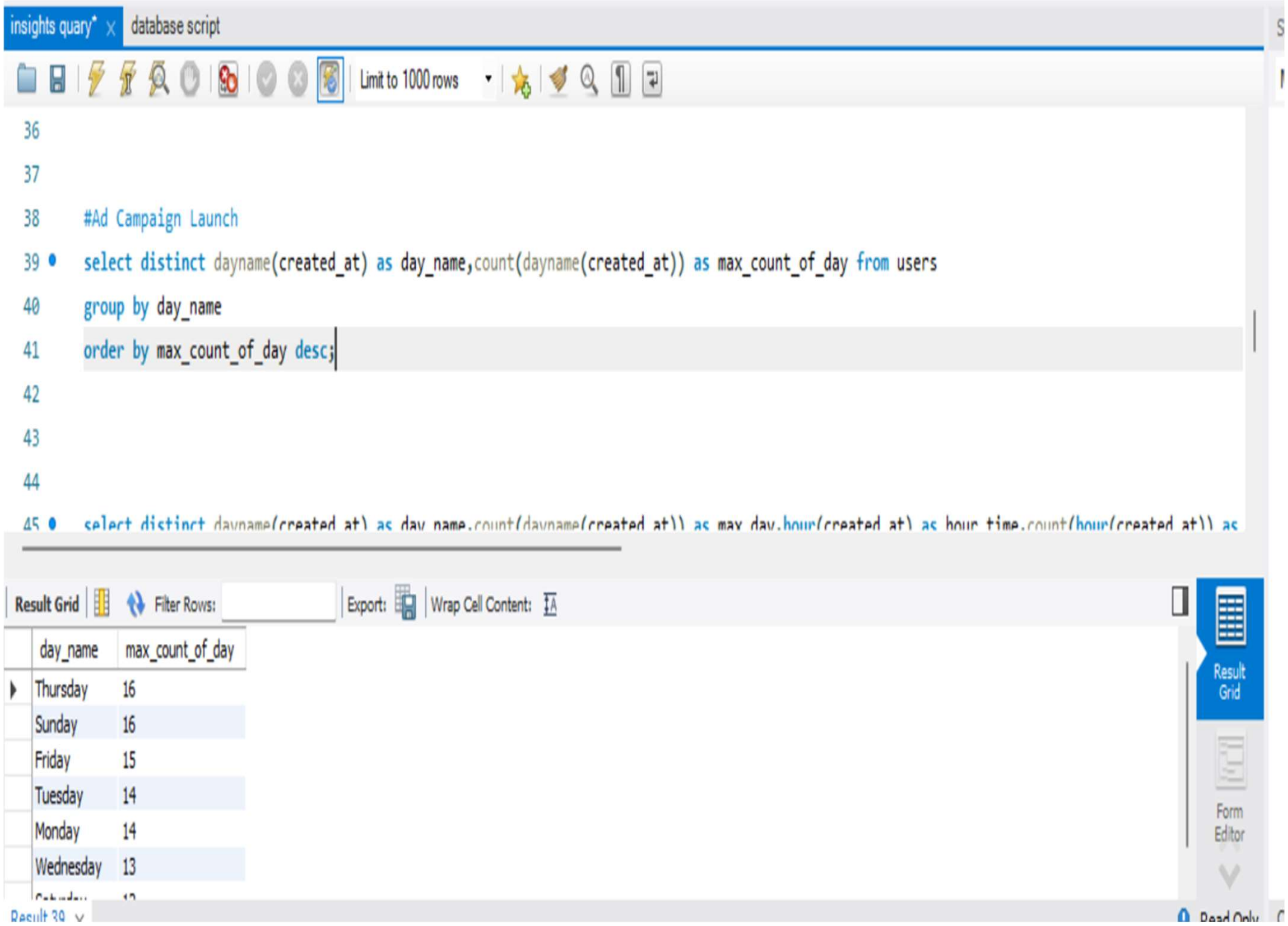
tag_name	max_tags
smile	59
beach	42
party	39
fun	38
concert	24

```
select tags.tag_name,count(tags.id) as max_tags from photo_tags
inner join tags on photo_tags.tag_id=tags.id
inner join photos on photo_tags.photo_id=photos.id
group by tag_name
order by max_tags desc
limit 5;
```

E. ***Ad Campaign Launch:*** *The team wants to know the best day of the week to launch ads.*

So ,I find out , in which day of a week most user are doing registration in Instagram.

****And after Deep Analysis I find That , in Thursday and Sunday max user's are doing registration and at the time between 19 p.m. to 20 p.m. of Sunday the no of registration is maximum.****



```
#Ad Campaign Launch
select distinct dayname(created_at) as day_name,count(dayname(created_at)) as max_count_of_day from users
group by day_name
order by max_count_of_day desc;



select distinct dayname(created_at) as day_name,count(dayname(created_at)) as max_day,hour(created_at) as hour_time,count(hour(created_at)) as
```

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

Result 39

select distinct dayname(created_at) as day_name,count(dayname(created_at)) as max_count_of_day from users

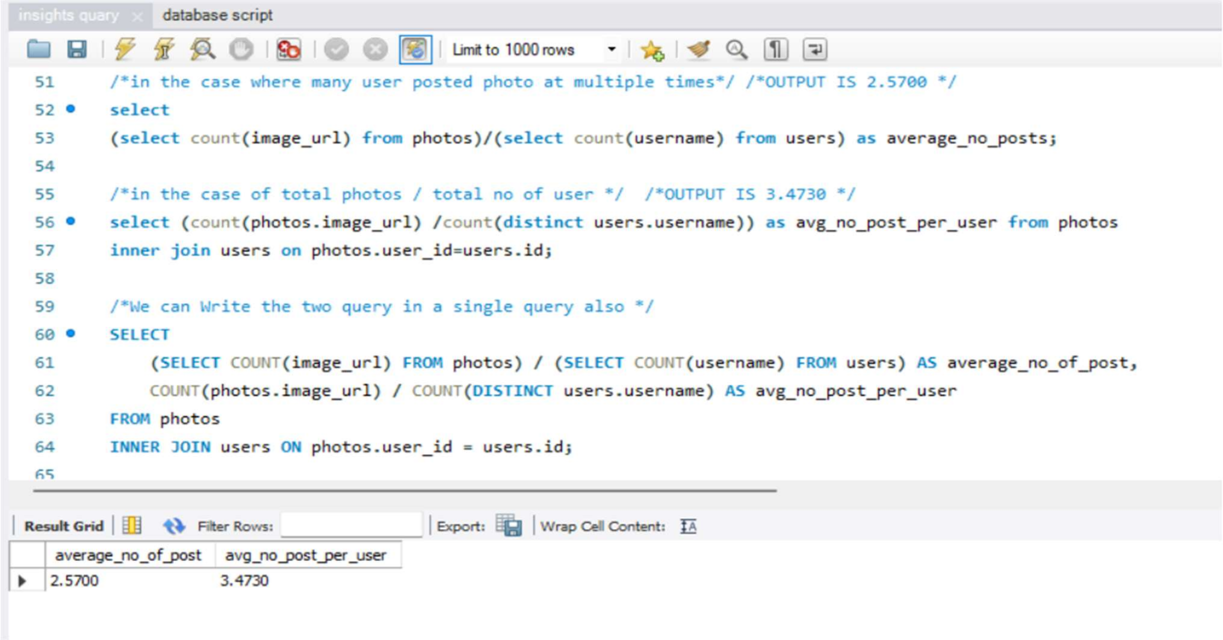
group by day_name

order by max_count_of_day desc;

F. ***User Engagement:*** *Investors want to know if users are still active and posting on Instagram or if they are making fewer posts.*

So, Here I calculated

1. *the average number of posts per user on Instagram(In the case ,where many user posted photo multiple times)*
2. *the total number of photos on Instagram divided by the total number of users*



```
/*in the case where many user posted photo at multiple times*/ /*OUTPUT IS 2.5700 */
select
(select count(image_url) from photos)/(select count(username) from users) as average_no_posts;

/*in the case of total photos / total no of user */  /*OUTPUT IS 3.4730 */
select (count(photos.image_url) /count(distinct users.username)) as avg_no_post_per_user from photos
inner join users on photos.user_id=users.id;

/*We can Write the two query in a single query also */
SELECT
    (SELECT COUNT(image_url) FROM photos) / (SELECT COUNT(username) FROM users) AS average_no_of_post,
    COUNT(photos.image_url) / COUNT(DISTINCT users.username) AS avg_no_post_per_user
FROM photos
INNER JOIN users ON photos.user_id = users.id;
```

average_no_of_post	avg_no_post_per_user
2.5700	3.4730

*/*in the case where many user posted photo at multiple times*/ /*OUTPUT IS 2.5700 */*

select

(select count(image_url) from photos)/(select count(username) from users) as average_no_posts;

*/*in the case of total photos / total no of user */ /*OUTPUT IS 3.4730 */*

select (count(photos.image_url) /count(distinct users.username)) as avg_no_post_per_user from photos

inner join users on photos.user_id=users.id;

*/*We can Write the two query in a single query also */*

SELECT

(SELECT COUNT(image_url) FROM photos) / (SELECT COUNT(username) FROM users) AS average_no_of_post,

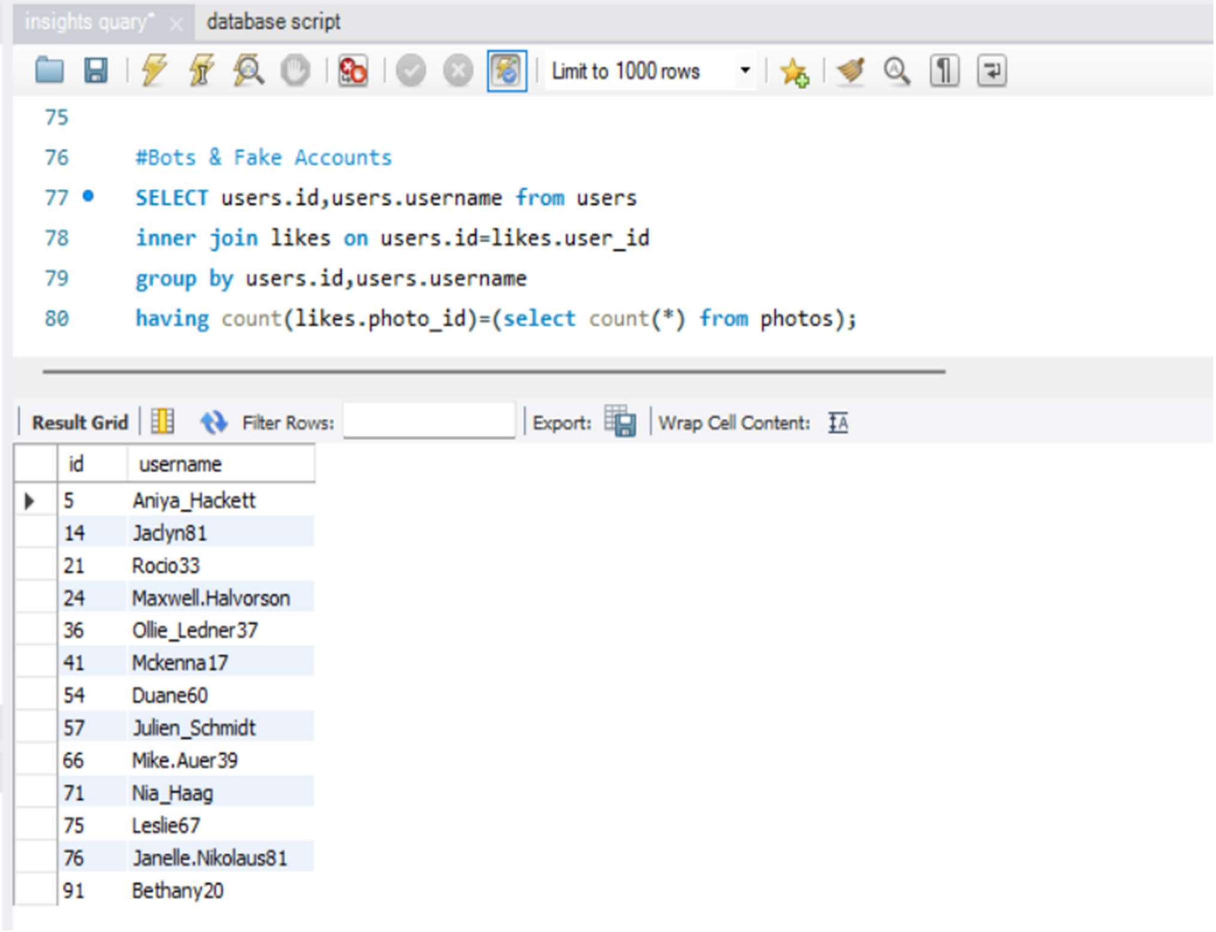
COUNT(photos.image_url) / COUNT(DISTINCT users.username) AS avg_no_post_per_user

FROM photos

INNER JOIN users ON photos.user_id = users.id;

G. ***Bots & Fake Accounts:*** *Investors want to know if the platform is crowded with fake and dummy accounts.*

So,to know this I extracted the user who are liked in all posts.



```
#Bots & Fake Accounts
SELECT users.id,users.username from users
inner join likes on users.id=likes.user_id
group by users.id,users.username
having count(likes.photo_id)=(select count(*) from photos);
```

id	username
5	Aniya_Hackett
14	Jaclyn81
21	Rocio33
24	Maxwell.Halvorson
36	Ollie_Ledner37
41	Mckenna17
54	Duane60
57	Julien_Schmidt
66	Mike.Auer39
71	Nia_Haag
75	Leslie67
76	Janelle.Nikolaus81
91	Bethany20

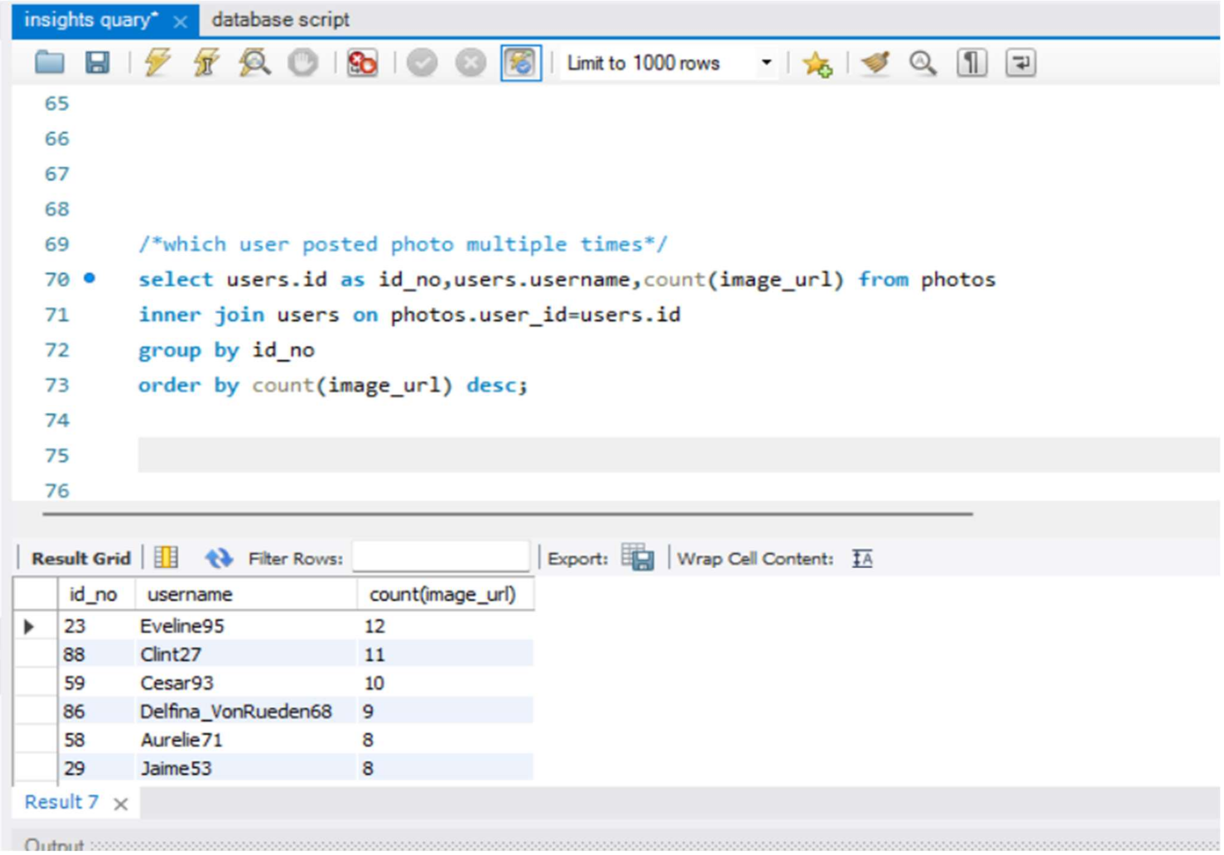
SELECT users.id,users.username from users

inner join likes on users.id=likes.user_id

group by users.id,users.username

having count(likes.photo_id)=(select count(*) from photos);

/ Along with This all I also find out the users who are posted photos multiple times*/*



```
/*which user posted photo multiple times*/
select users.id as id_no,users.username,count(image_url) from photos
inner join users on photos.user_id=users.id
group by id_no
order by count(image_url) desc;
```

id_no	username	count(image_url)
23	Eveline95	12
88	Clint27	11
59	Cesar93	10
86	Delfina_VonRueden68	9
58	Aurelie71	8
29	Jaime53	8

select users.id as id_no,users.username,count(image_url) from photos

inner join users on photos.user_id=users.id

group by id_no

order by count(image_url) desc;

5.Results → So, I have gained practical knowledge of Data Analysis. Also, I learned work flow of Data Analytics. I understand that Business understanding is also very important for Data Analysis and realized that playing with Data is not too easy we have to understand the Data first very well, after that we will be able to get some important insights from our Data. After doing this project I have improved my showcasing skill also.

So, I have find out the above insights which will be help to take decision and grow the business in this competitive market.