Instagram User Analytics Report

Project Description

The Project is “Instagram User Analytics” where we try to understand the users of Instagram app and their interactions to gain insights about users. The main goal of analysis is to find Insights so that we can understand users, what can be improved like adding new features, promoting ad campaigns based on user interest, what needs to be done to get more users interactions.

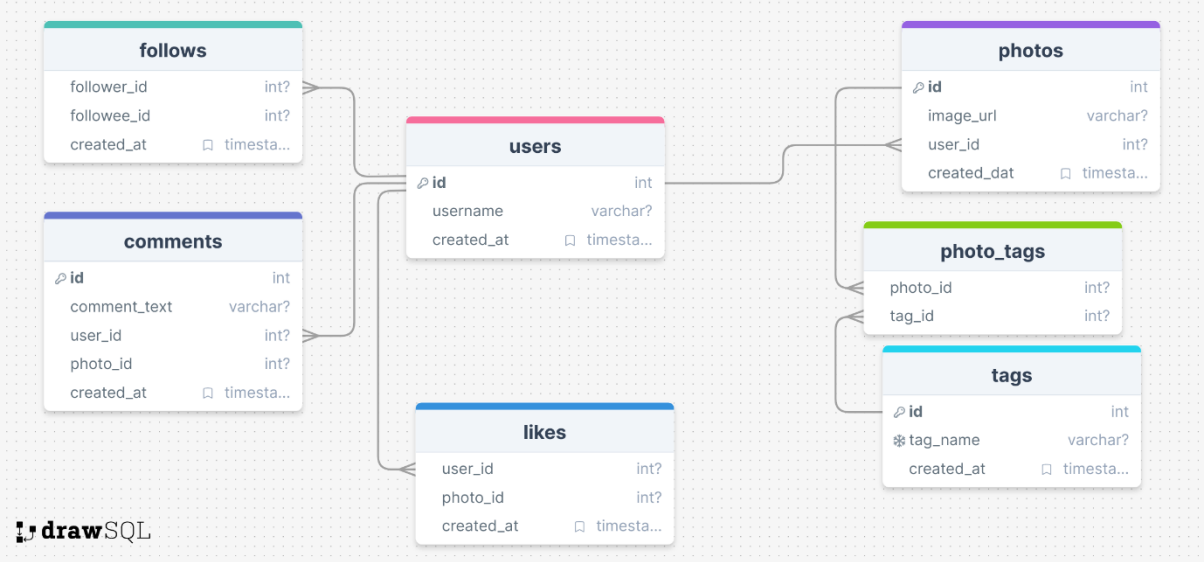
All of these Insights are addressed in this Project by answering few of sql questions. In this project I have used MySQL to extract required data from dataset and answer the questions through data.

Approach

Dataset Creation and Understanding:

I have created the tables in the database and named the database as ig_clone.

I have tried to understand the dataset given to me. The relationship between different tables and finally made a sql schema diagram for better understanding which is as follows.



In this schema,

- 🔑 symbol – Primary key of table
- ? symbol – Not Null
- ❄ symbol – Unique value
- 🔖 symbol – column attributes (set default time as now())

The dataset describes about the data of Instagram users from May 20216 to May 2017 which is about a year's data.

Steps used to find the Insights:

I have analyzed the datasets and answered few questions to get useful insights. The details of answered questions are,

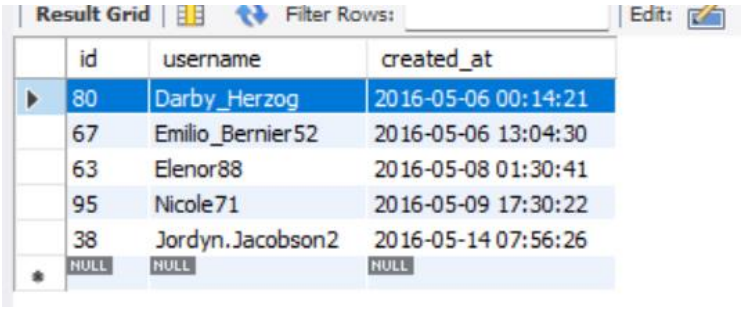
Marketing Analysis questions:

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

   SQL query :

   select * from users order by created_at limit 5;

   output:

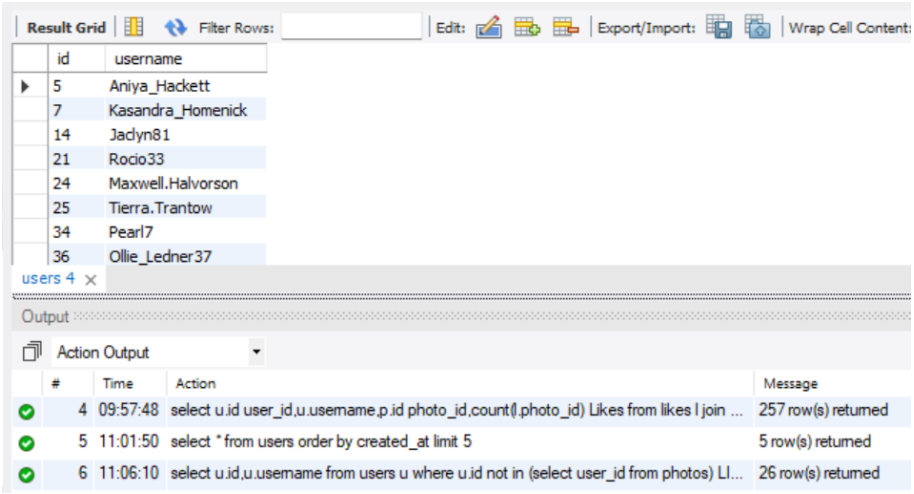


2. Identify users who have never posted a single photo on Instagram.

   SQL query:

   select u.id,u.username from users u where u.id not in (select user_id from photos);

   output:



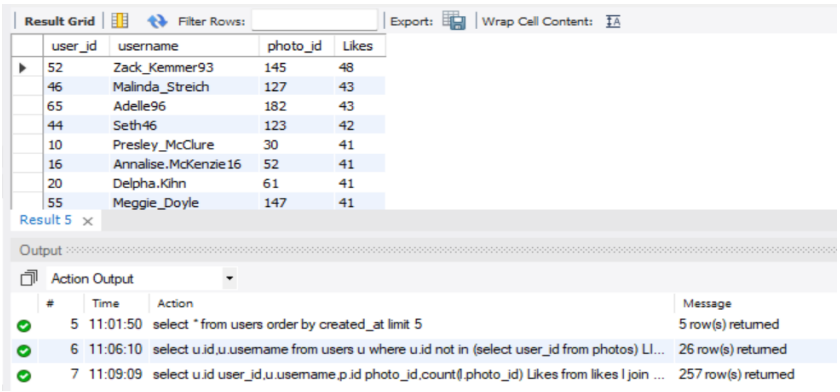
3. Contest Winner Declaration: The team has organized a contest where the user with the most likes on a single photo wins.

Your Task: Determine the winner of the contest and provide their details to the team.

SQL query :

select u.id user_id,u.username,p.id photo_id,count(l.photo_id) Likes from likes l join photos p on p.id=l.photo_id join users u on u.id=p.user_id group by photo_id order by Likes desc;

output:



user_id	username	photo_id	Likes
52	Zack_Kemmer93	145	48
46	Malinda_Streich	127	43
65	Adelle96	182	43
44	Seth46	123	42
10	Presley_McClure	30	41
16	Annalise.McKenzie16	52	41
20	Delpha.Kihn	61	41
55	Meggie_Doyle	147	41

Result 5

Output

Action Output

#	Time	Action	Message
5	11:01:50	select * from users order by created_at limit 5	5 row(s) returned
6	11:06:10	select u.id,u.username from users u where u.id not in (select user_id from photos) LI...	26 row(s) returned
7	11:09:09	select u.id user_id,u.username,p.id photo_id,count(l.photo_id) Likes from likes l join ...	257 row(s) returned

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

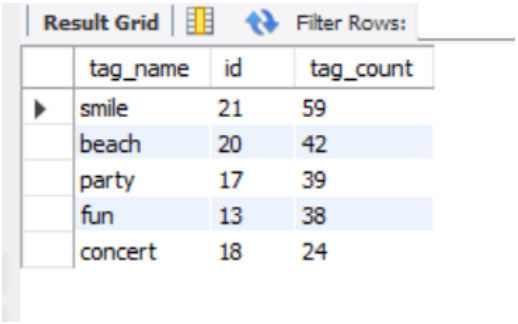
SQL query :

select t.tag_name,t.id,count(*) tag_count from photo_tags p join tags t on p.tag_id=t.id

group by p.tag_id

order by tag_count desc limit 5;

output:



tag_name	id	tag_count
smile	21	59
beach	20	42
party	17	39
fun	13	38
concert	18	24

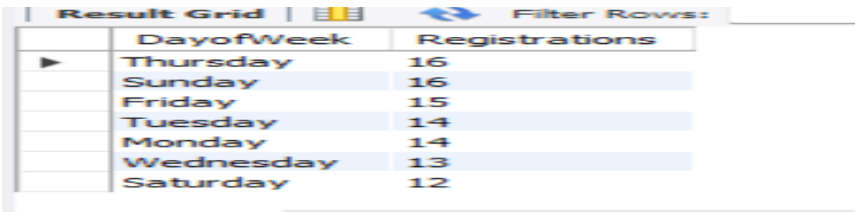
5. Determine the day of the week when most users register on Instagram.

SQL query :

select dayname(created_at) DayofWeek,count(created_at) Registrations from users

group by DayofWeek order by Registrations desc;

output:



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

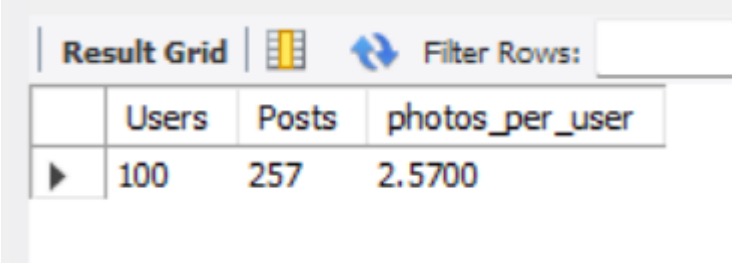
Investor Metrics Questions:

1. 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.

SQL query :

```
select count(u.id) Users,(select count(p1.id) from photos p1) Posts,
(select count(id) from photos)/count(u.id) photos_per_user from users u;
```

output:



Result Grid | Filter Rows:

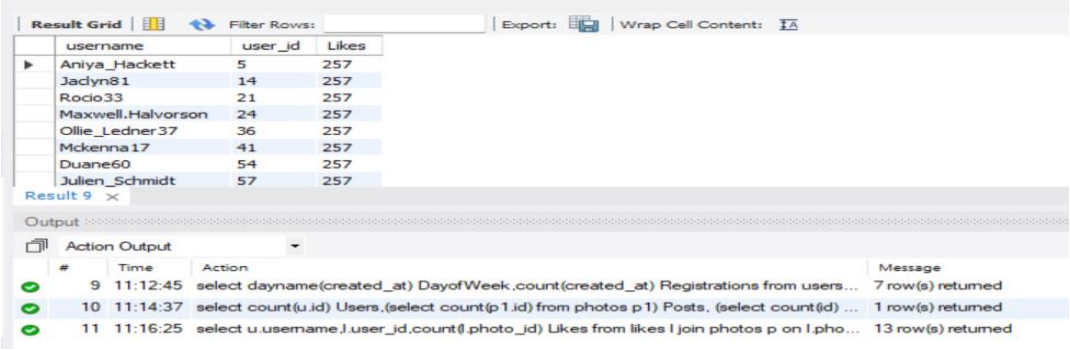
	Users	Posts	photos_per_user
▶	100	257	2.5700

2. Identify users (potential bots) who have liked every single photo on the site, as this is not typically possible for a normal user.

SQL query :

```
select u.username,l.user_id,count(l.photo_id) Likes from likes l join photos p on l.photo_id=p.id
join users u on l.user_id=u.id
group by l.user_id having Likes = (select count(p.id) from photos p);
```

output :



Result Grid | Filter Rows: | Export: | Wrap Cell Content:

	username	user_id	Likes
▶	Aniya_Hackett	5	257
	Jadyn81	14	257
	Rocio33	21	257
	Maxwell.Halvorson	24	257
	Ollie_Ledner37	36	257
	Mckenna17	41	257
	Duane60	54	257
	Julien_Schmidt	57	257

Result 9

Output

Action Output

#	Time	Action	Message
9	11:12:45	select dayname(created_at) DayofWeek,count(created_at) Registrations from users...	7 row(s) returned
10	11:14:37	select count(u.id) Users,(select count(p1.id) from photos p1) Posts, (select count(id) ...	1 row(s) returned
11	11:16:25	select u.username,l.user_id,count(l.photo_id) Likes from likes l join photos p on l.pho...	13 row(s) returned

Tech-Stack Used

The Tech-Stack I have used are as follows,

- MySQL Workbench
    - MySQL workbench is best tool to use when dealing with Structured data using MySQL.
    - The Interface is User Friendly to use.

Insights

- About 26% of users never created any posts among all other users.
- There are a good number of users who are showing interest in participating in contests conducted. Contests can give some sort of award to gain more participants.
- Top 5 tags are smile, beach, party, fun, concert Which says that users are mostly interested in Entertainment related content so it is best to collaborate with brands who are Promoting Entertainment and related fields.
- The partner brands can create posts related to these 5 hashtags to reach most people.
- Most users are registering on Thursdays and Sundays so that would be good time to do an ad campaign.
- There are 13 bots among 100 users which shows that most of the users are real humans and not bots. So, Investors can invest in the app.

Result

I have gained a lot of knowledge by doing this project like

- Creating, Understanding and Cleaning of datasets.
- Extracting useful insights from the datasets which can help in the growth of Instagram app.
- By writing this report I was able to describe my findings in a clear and clean format.

The Insights that I have found out are useful for the Instagram app to gain more Users, make more Partners and Collaborations or even to get investors to invest in their app.