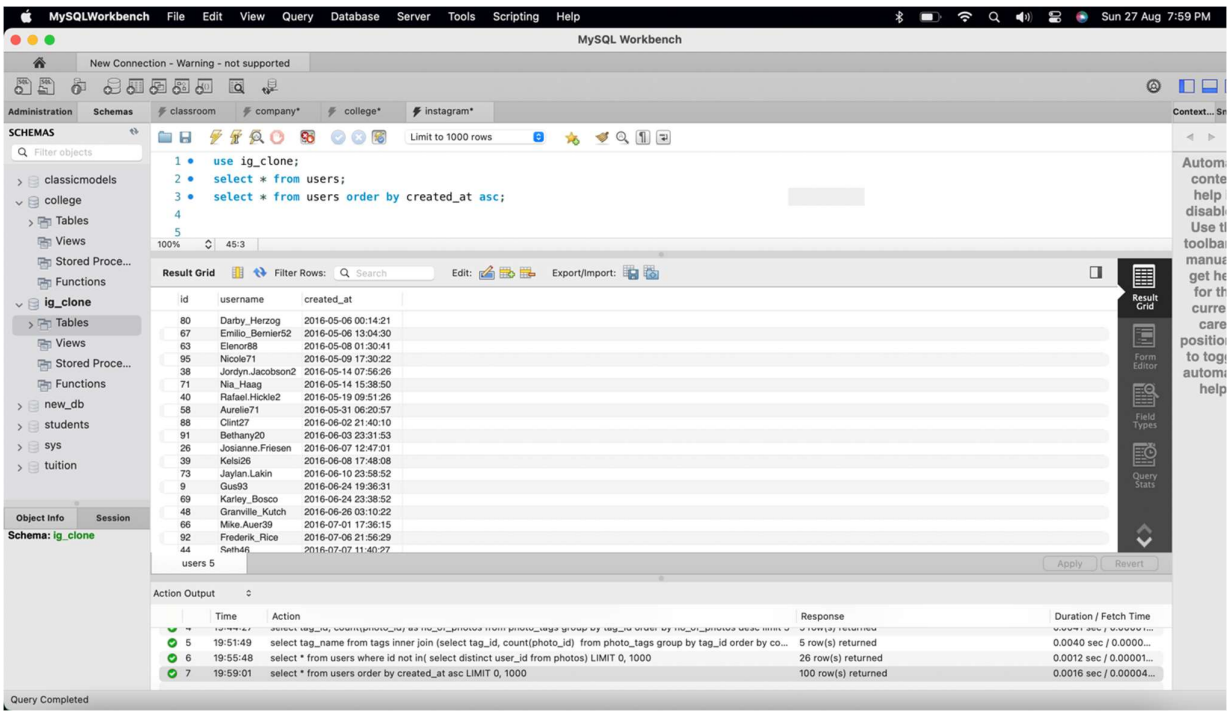
Project Description: The main purpose of this project is to find those data from Instagram through which we can find regular, real user accounts to give them incentives and to find those users who seems to be less active should be motivated through some lucrative ideas and if they are not active for a long period of time, then the account should be deactivated. In short our main focus will be company's growth.

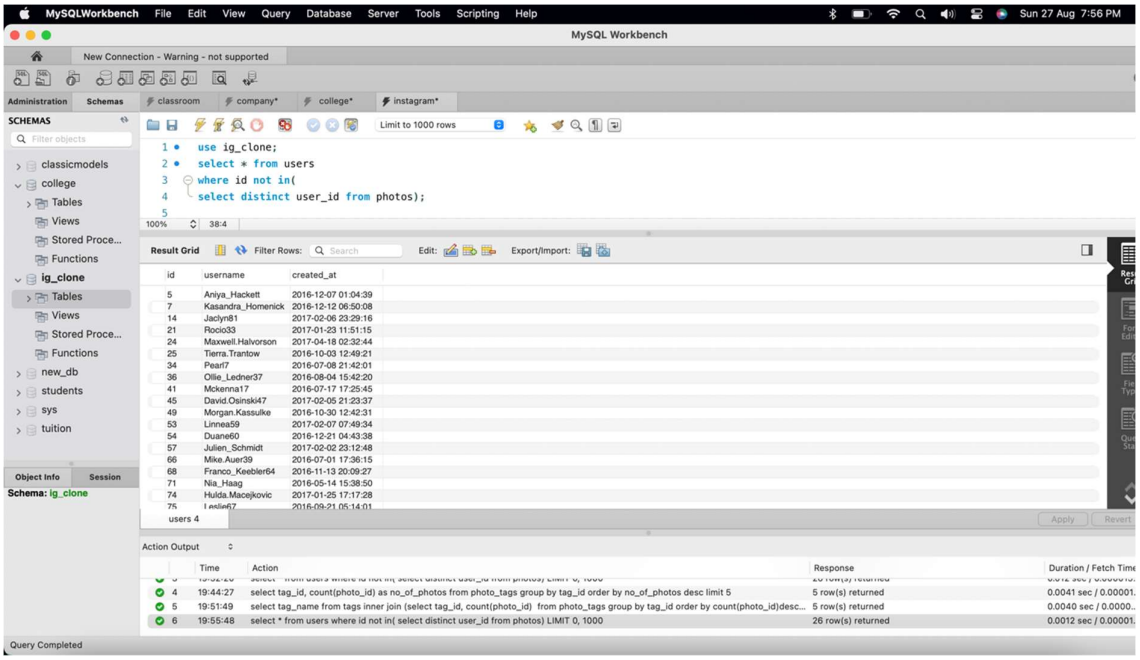
The approach we'll take to handle this project by collecting data in our software and will perform queries on them according to our target and find the result to derive the insights.

Approach:

1) Performed order by on created_at column column in users table.



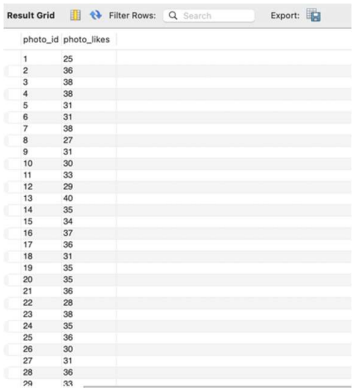
2) Selected those user id which were not present in photos column through subqueries.



3) First, I viewed the likes table then I noticed that by grouping that data of photo id with counting of user id we can find the number of likes on a single photo and can order them accordingly. So by grouping the data I made a table.

```
use ig_clone;
select * from likes;
select photo_id, count(user_id) as photo_likes
from likes
group by photo_id;
```

Using these queries , I grouped the table



Result Grid Filter Rows: Search Export:

photo_id	photo_likes
1	25
2	36
3	38
4	38
5	31
6	31
7	38
8	27
9	31
10	30
11	33
12	29
13	40
14	35
15	34
16	37
17	36
18	31
19	35
20	35
21	36
22	28
23	38
24	35
25	36
26	30
27	31
28	36
29	33

Then I found the maximum value of photo_likes using subqueries

```
select max(photo_likes)
from (select photo_id, count(user_id) as photo_likes
from likes
group by photo_id)
as MaxphotoLikes;
```

100% 18:6

Result Grid Filter Rows: Search Export:

max(photo_lik...
48

After finding the maximum value we found the photo id according to that by using that group by from table and using that maximum values query in where clause with photo_likes so that we can find the photo_id where the likes are maximum. Following query we used.

```
select photo_id
from (select photo_id, count(user_id) as photo_likes
from likes
group by photo_id
) as total_likes
where photo_likes = (select max(photo_likes)
from (select photo_id, count(user_id) as photo_likes
from likes
group by photo_id
) as total_likes);
```

100% 19:16

Result Grid Filter Rows: Search Export:

photo_id
145

Now we are going to use this query as subquery in another query to find the user id of that person who posted that photo. So, we just copied the previous command of photo_id in where clause for finding the users id of that photo id in users table.

```
select user_id from photos
where id= ( select photo_id
from (select photo_id, count(user_id) as photo_likes
from likes
group by photo_id
) as total_likes
where photo_likes = (select max(photo_likes)
from (select photo_id, count(user_id) as photo_likes
from likes
group by photo_id
) as total_likes));
```

100% 20:28

Result Grid Filter Rows: Search Export:

user_id
52

Now, using this user_id we will print the users information by copying that query in users table as a subquery.

In where clause of id of pasted the query through which we found the user id.

```
select * from users
where id = (select user_id from photos
where id = (select photo_id
from(select photo_id,count(user_id) as photo_likes
from likes
group by photo_id
) as total_likes
where  photo_likes = (select max(photo_likes)
from(select photo_id,count(user_id) as photo_likes
from likes
group by photo_id
) as total_likes
)));
```

0% 5:14

Result Grid Filter Rows: Search Edit: Export/Import:

id	username	created_at
52	Zack_Kemmer93	2017-01-01 05:58:22
NULL	NULL	NULL

users 8

4) In the photo_tag table data I found that we can count the number of photo and group by it with tag id to find the max tag used. Through this approach we made a query which prints the top 5 tags used through order by and limit :

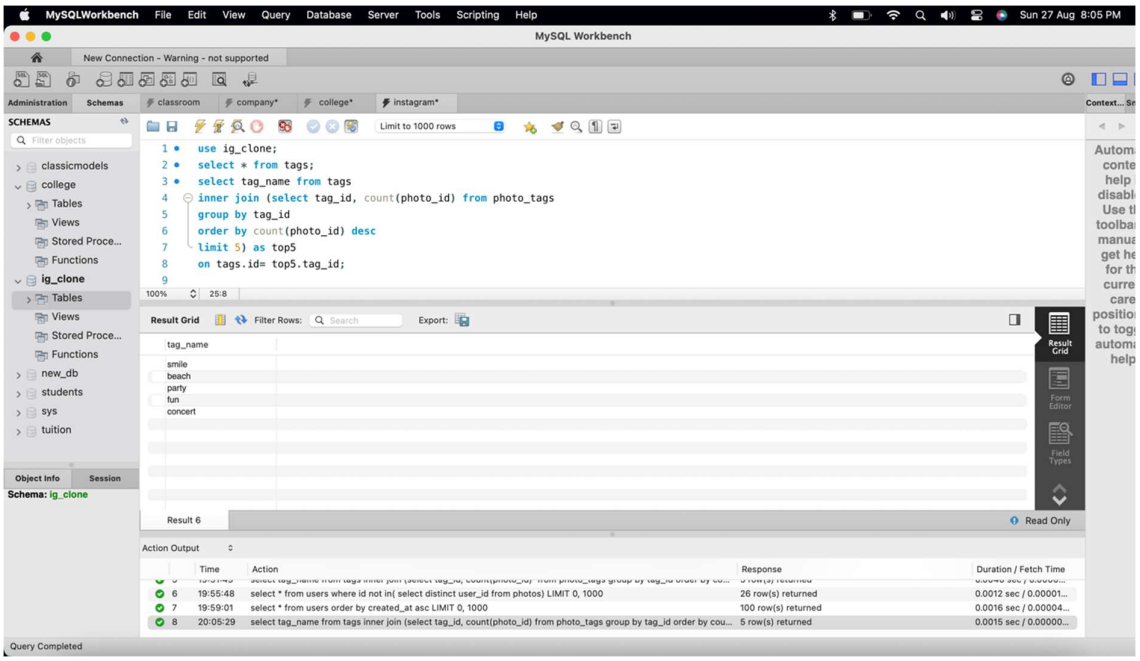
```
select tag_id, count(photo_id) as no_of_photos
from photo_tags
group by tag_id
order by no_of_photos desc
limit 5;
```

0% 9:6

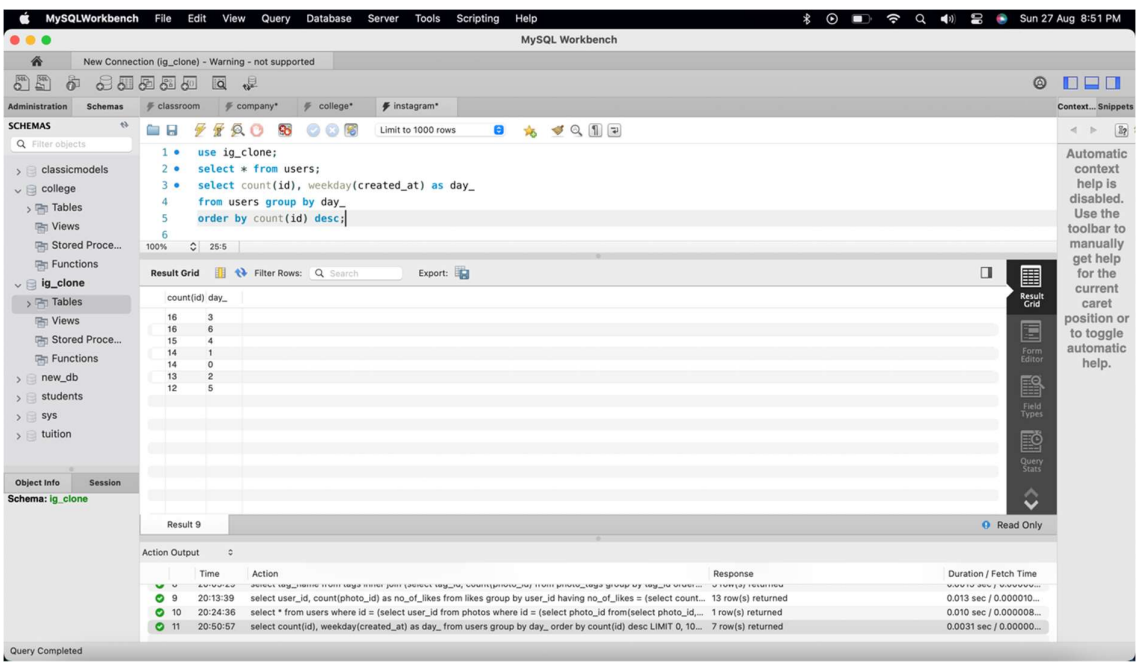
Result Grid Filter Rows: Search Export: Fetch rows:

tag_id	no_of_phot...
21	59
20	42
17	39
13	38
18	24

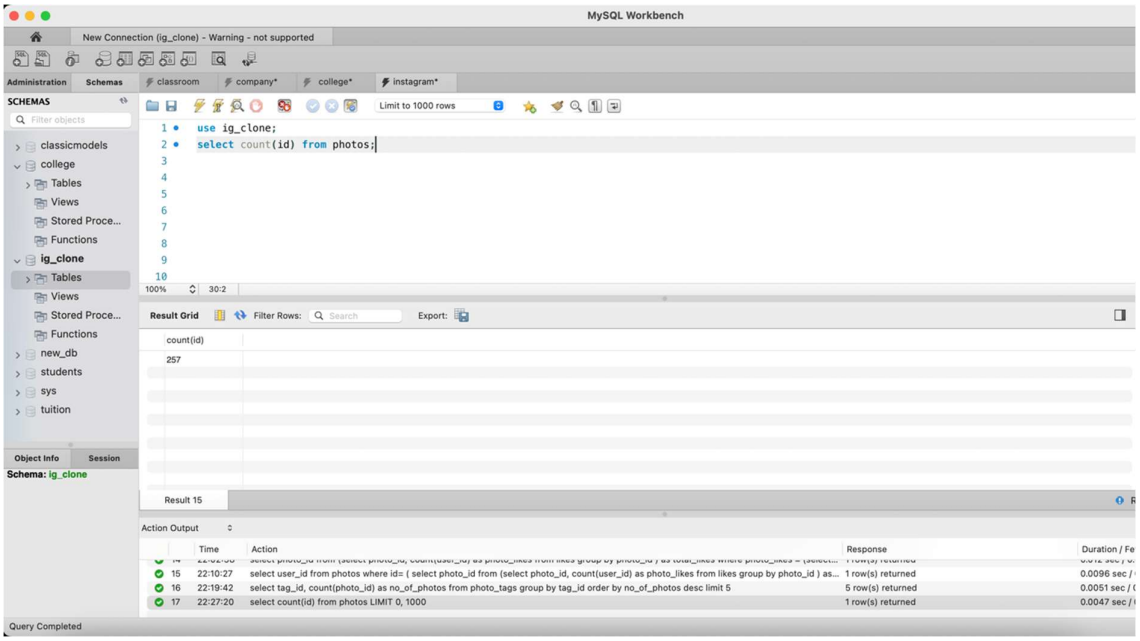
And then to find the tag name of that id we inner joined the table with tags table



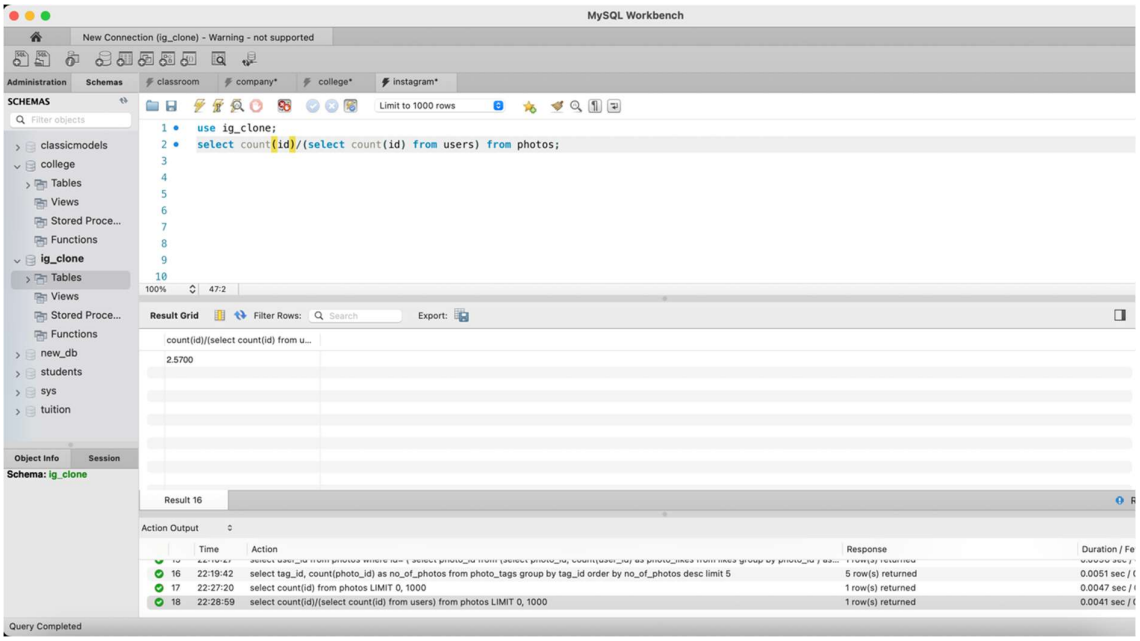
5) I grouped the data of day with count of users id that were there for that day.



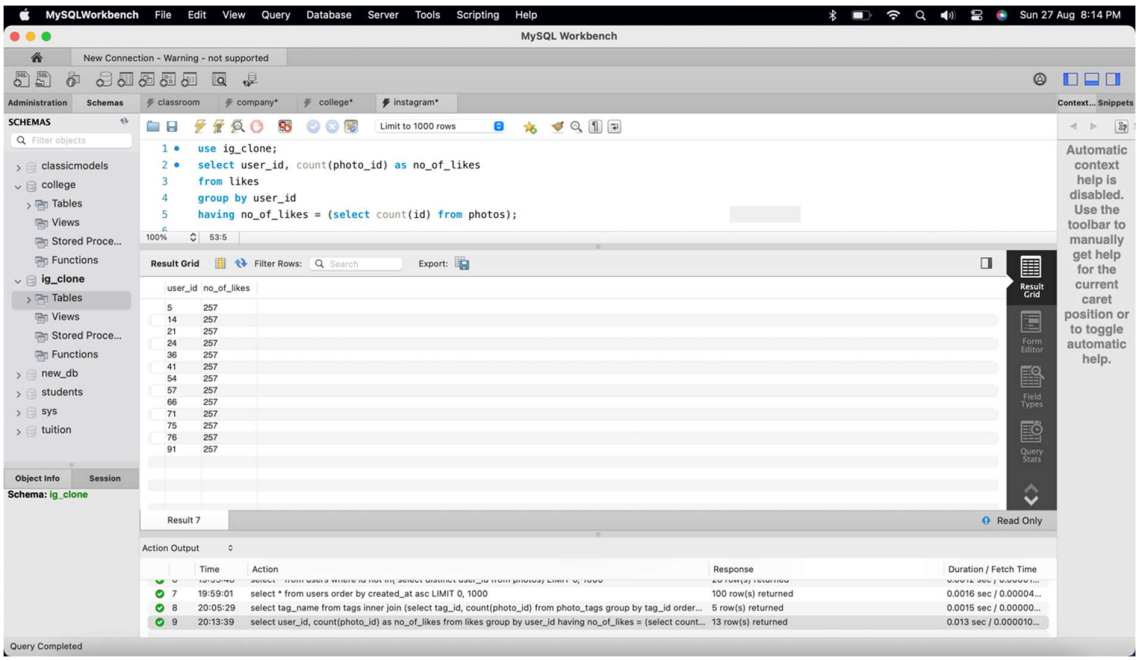
6) Found total number of photos by counting it.



Dividing it by number of id's.



7) Grouped the data of user_id with count of photo_id having those rows where number of likes is equal to the total number of photos which gives those users who liked every photo.

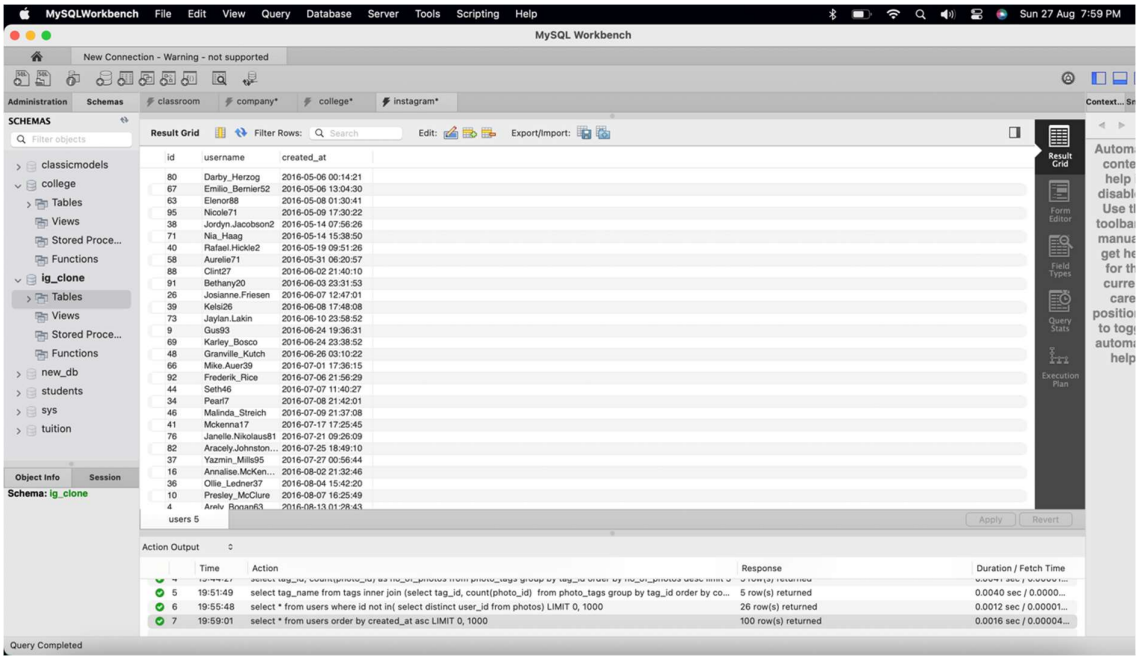


Tech-Stack Used: MySQL Workbench.

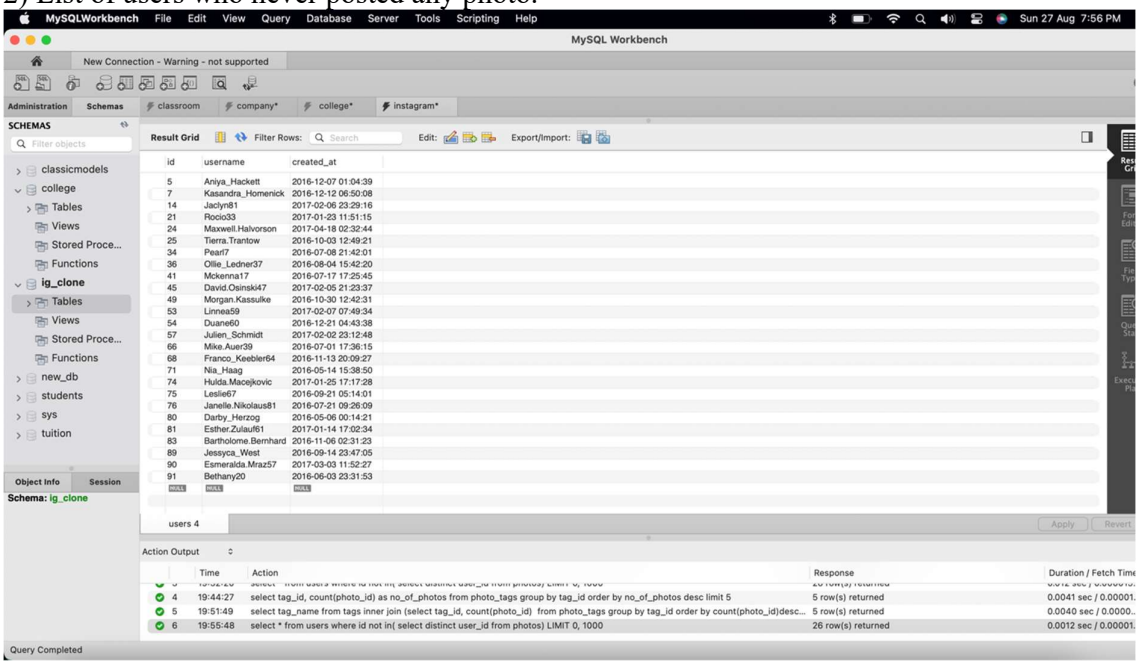
My SQL Workbench is an easy to use tool for SQL through which we can perfrom different queries on database and work on them

Insights:

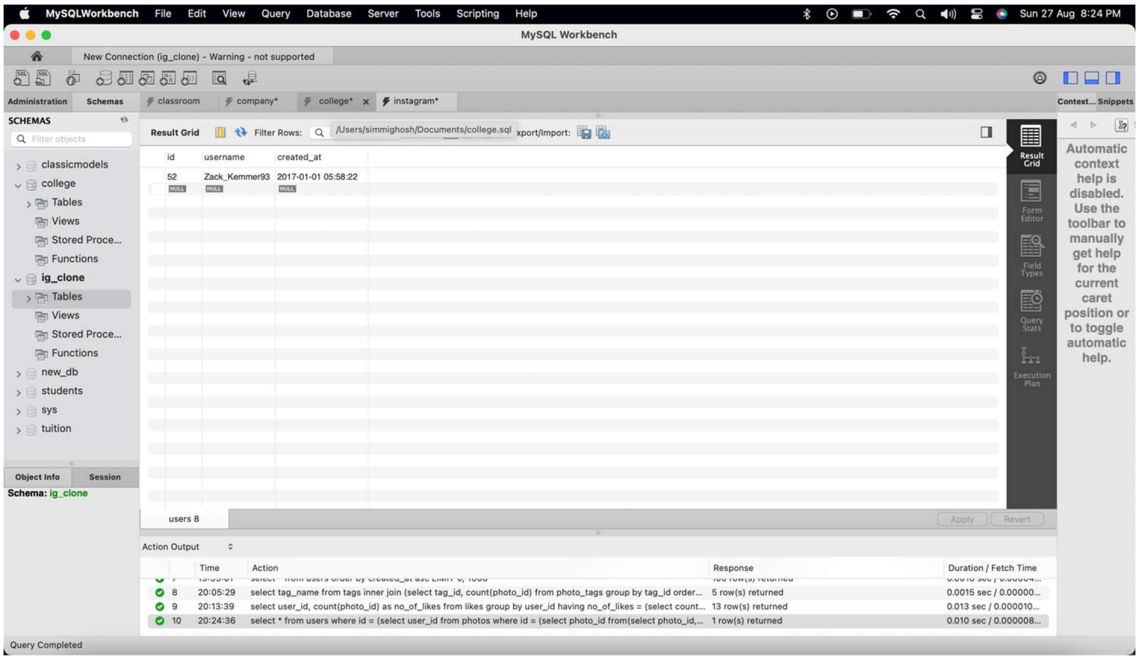
1)List of users on Instagram according to joining data. User joined first is on top.



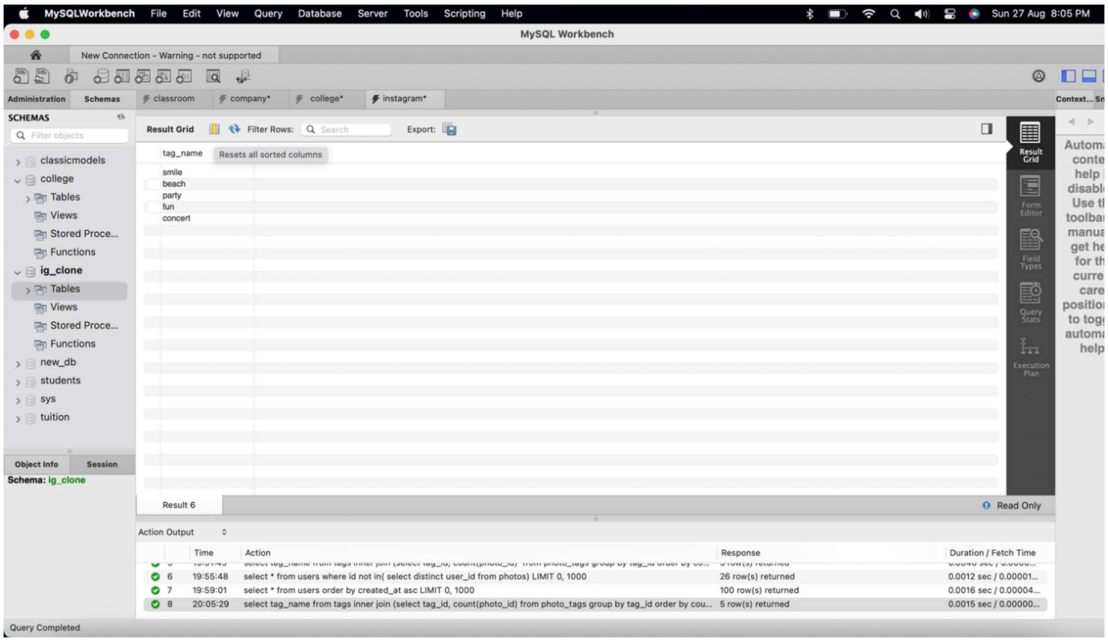
2) List of users who never posted any photo.



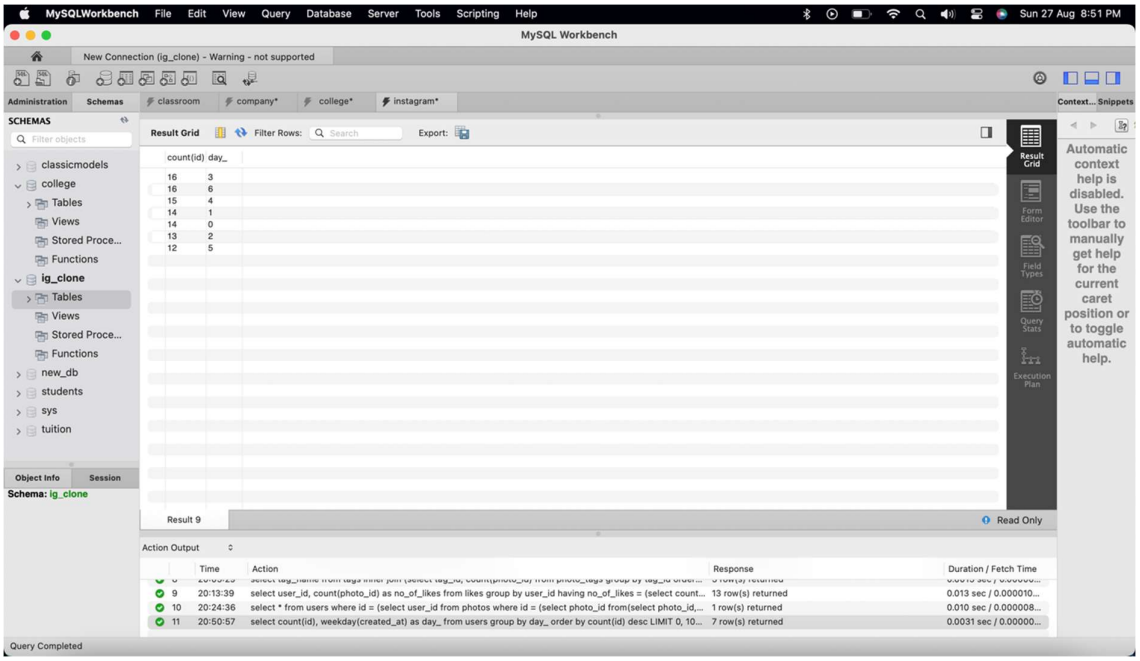
3)Winner of contest with 48 likes on a single photo.



4) List of top 5 trending tags.

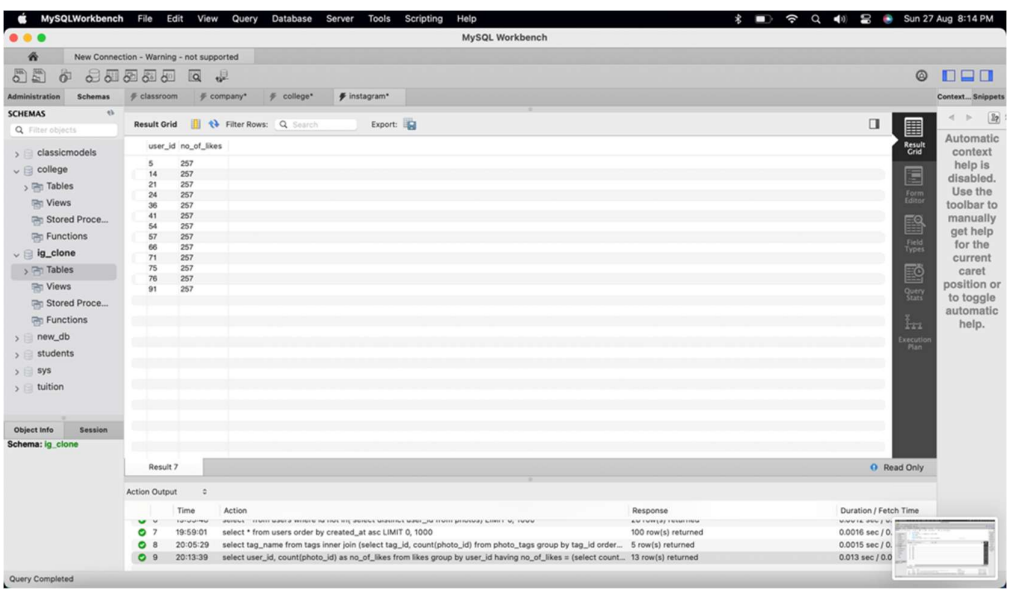


5) List of users joining according to the day. 0 means Monday.



6) The total number of photos is 257 and total number of users is 100. Average of photos per users is 2.57.

7) List of people who liked every single photo.



Result:

A) Market Analysis:

1. The most loyal or five oldest users of Instagram are:

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

2. The users who have never posted any photo are:

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
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. The winner of contest of most likes on a single photo is:

id	username	created_at
52	Zack_Kemmer93	2017-01-01 05:58:22

4. The most popular hashtags are:

tag_name
smile
beach
party
fun
concert

Resets all sorted columns

5. The best days to launch ads on Instagram are Wednesdays and Sundays as most user registers on Wednesdays and Sundays. So, we can schedule our ad campaign on next Wednesday and Sunday that is 30th august and 3rd of September.

B) Investor Metrics:

1. The average of posts per user on Instagram is 2.57
And total number of photos on Instagram are 257.

2. The list of bots and fake accounts is given below:

user_id	no_of_likes
5	257
14	257
21	257
24	257
36	257
41	257
54	257
57	257
66	257
71	257
75	257
76	257
91	257