Name: Shraddha Deori

Instagram User Analysis

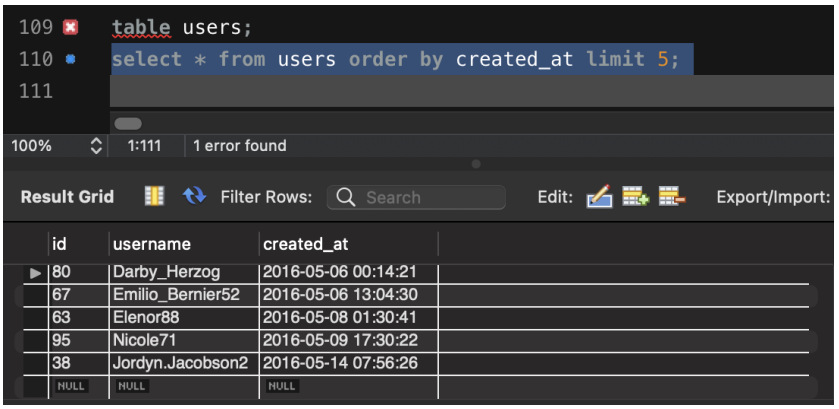
The first thing I did was to search videos in youtube giving informations on how to apply queries once more. The queries were a little advanced, so it took sometime to complete, but it got completed at last. I searched up using mysql data in jupyter notebook, but got a little confused, so that is a topic I need to get my hands on again.

I learned how to find out answers to the queries which are generally asked in a company, for eg, the total number of accounts which are real and not bots, the number of times each user has posted anything, or the trending hashtags in a particular website.

The queries given were very practical and are used on a daily basis in companies which is a very useful practice to students like us, who are beginners. I learned a lot of things.

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



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
NULL	NULL	NULL

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

```
109  table users;
110  table photos;
111
112  select users.username, users.id, photos.user_id
113  from users
114  LEFT JOIN photos
115  ON users.id = photos.user_id
116  WHERE photos.user_id is null;
117
```

username	id	user_id
Aniya_Hackett	5	NULL
Kasandra_Homenick	7	NULL
Jaclyn81	14	NULL
Rocio33	21	NULL
Maxwell.Halvorson	24	NULL
Tierra.Trantow	25	NULL
Pearl7	34	NULL
Ollie_Ledner37	36	NULL
Mckenna17	41	NULL
David.Osinski47	45	NULL
Morgan.Kassulke	49	NULL
Linnea59	53	NULL
Duane60	54	NULL
Julien_Schmidt	57	NULL
Mike.Auer39	66	NULL
Franco_Keebler64	68	NULL
Nia_Haag	71	NULL
Hulda.Macejkovic	74	NULL
Leslie67	75	NULL
Janelle.Nikolaus81	76	NULL
Darby_Herzog	80	NULL
Esther.Zulauf61	81	NULL
Bartholome.Bernhard	83	NULL
Jessyca_West	89	NULL
Esmeralda.Mraz57	90	NULL
Bethany20	91	NULL

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

```
109  table users;
110  table photos;
111  table likes;
112  table tags;
113
114  select username, photos.id, photos.image_url, count(*) as total
115  from photos
116  inner join likes
117  on likes.photo_id = photos.id
118  inner join users
119  on photos.user_id = users.id
120  group by photos.id
121  order by total desc
122  limit 1;
123
```

100% 1:113

Result Grid | Filter Rows: Search | Export: | Fetch rows:

username	id	image_url	total
Zack_Kemmer93	145	https://jarret.name	48

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

```
108
    table users;
110 table photos;
111 table likes;
112 table tags;
113
114 select tags.tag_name, count(*) as total
115 from photo_tags
116 join tags
117 on photo_tags.tag_id = tags.id
118 group by tags.id
119 order by total
120 limit 5;
121
```

100% 1:122

Result Grid | Filter Rows: Search | Export: | Fetch rows:

tag_name	total
foodie	11
delicious	15
stunning	16
photography	16
sunrise	17

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

```
109 table users;
110 table photos;
111 table likes;
112 table tags;
113
114 select dayname(created_at) as day, count(*) as total
115 from users
116 group by day
117 order by total desc
118 limit 7;
119
```

100% 1:113

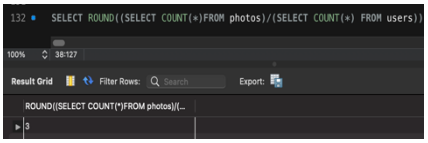
Result Grid | Filter Rows: Search | Export: | Fetch rows:

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

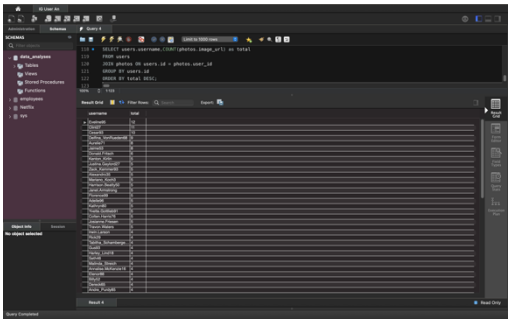
B) Investor Metrics: Our investors want to know if Instagram is performing well and is not becoming redundant like Facebook, they want to assess the app on the following grounds

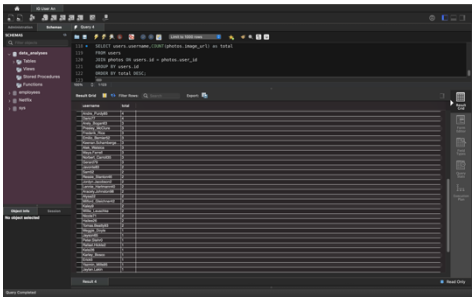
1. **User Engagement:** Are users still as active and post on Instagram or they are making fewer posts
   Your Task: Provide how many times does average user posts on Instagram. Also, provide the total number of photos on Instagram/total number of users

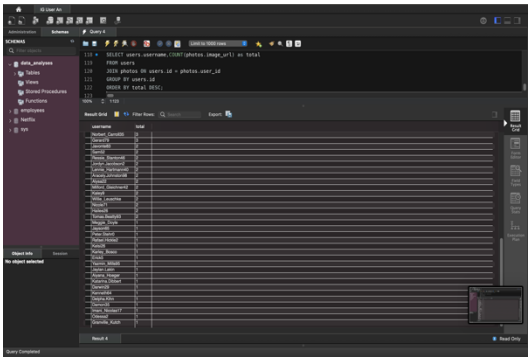
   Average no. of posts per user:



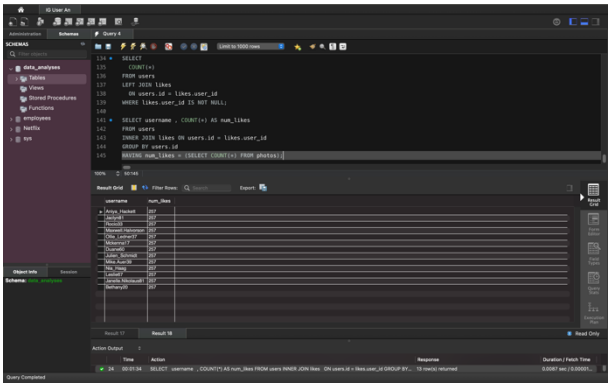
Total no. of photos per users:







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).


SELECT
COUNT(*)
FROM users
LEFT JOIN likes
ON users.id = likes.user_id
WHERE likes.user_id IS NOT NULL;
SELECT username , COUNT(*) AS num_likes
FROM users
INNER JOIN likes ON users.id = likes.user_id
GROUP BY users.id
DATING num_likes = (SELECT COUNT(*) FROM photos);
Result Grid
Action Output
Query Completed