Project Description: In this project we are trying to analyse instagram dataset. We will try to draw some key insights which can be used to grow the business (like finding bot users)

Approach: At first created the database ig_clone, Then created tables like users,photos,likes,follows,comments,tags,photo_tags. Then inserted the data into the respective tables. Then executed some SQL queries as per need to get insight from the data.

Tech-Stack Used: MySQL Server as the database, MySQL WORKBENCH as a query editor **Insights:** Below insights we could make like:

- 5 oldest users
- Users who have not yet uploaded a single photo
- The user whose photo got the maximum like and the photo details
- Top 5 commonly used hashtags
- On which day maximum users were registered on Instagram
- Total users , Total photos, avg photo per user
- Photos per User
- Bot Users on instagram

SQL Queries & Outputs:

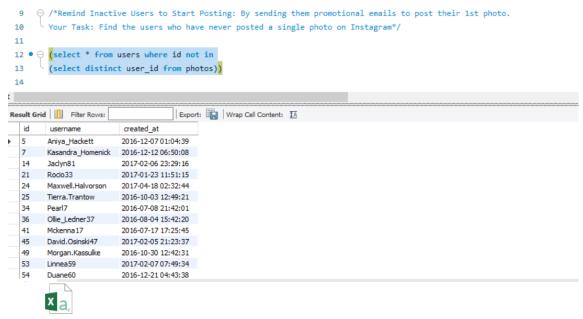
use ig clone;

/* 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*/

```
select * from users
order by created_at asc
limit 5;
      ⊖ /* 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*/
  3
  4
         select * from users
         order by created_at asc
  6
  7
         limit 5;
Edit: 🚰 🖶 Export/Import: 📳 🌄 Wrap Cell Content: 🏗 Fetch rows:
    id
         username
                         created_at
                        2016-05-06 00:14:21
   80
         Darby_Herzog
   67 Emilio_Bernier52 2016-05-06 13:04:30
                        2016-05-08 01:30:41
   63
        Elenor88
   95
        Nicole 71
                        2016-05-09 17:30:22
         Jordyn. Jacobson 2 2016-05-14 07:56:26
```

/*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*/

(select * from users where id not in (select distinct user_id from photos))



Users who have not posted single photo.c

/*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 */

#with cte approach

```
with cte1 as
(select photo_id,count(*) as total_likes
from likes
group by photo_id
order by 2 desc
limit 1),
cte2 as
(select *
from photos p
join cte1 c
on p.id=c.photo_id )
select user_id,photo_id,image_url,total_likes from cte2
```

```
21 With ctel as
22 (select photo_id,count(*) as total_likes
       from likes
23
24
       group by photo_id
       order by 2 desc
25
26
      limit 1),
       cte2 as
27
from photos p
29
30
       join cte1 c
      on p.id=c.photo_id )
31
32
     select u.id,u.username,c2.photo_id,c2.image_url,c2.total_likes
       from users u join cte2 c2
33
       on u.id=c2.user_id
Result Grid Filter Rows:
                                 Export: Wrap Cell Content: TA
       username
                    photo_id image_url
                                           total_likes
       Zack_Kemmer93 145
                            https://jarret.name 48
```

#Another Way

select u.username,p.id,p.image_url,COUNT(*) AS total_likes
from photos p
join likes l
on l.photo_id = p.id
join users u
on p.user_id = u.id
group by p.id
order by 4 desc
limit 1;

/* Hashtag Researching: A partner brand wants to know, which hashtags to use in the post to reach the most people on the platform.

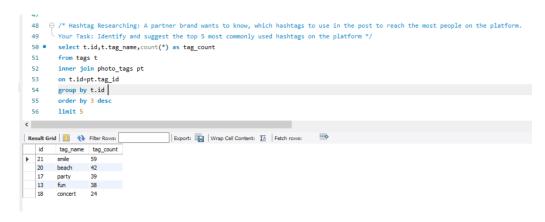
Your Task: Identify and suggest the top 5 most commonly used hashtags on the platform */ select t.id,t.tag_name,count(*) as tag_count

from tags t inner join photo_tags pt on t.id=pt.tag_id group by t.id order by 3 desc limit 5

#Another Way

with cte1 as(
select tag_id,count(*) as tag_count
from photo_tags
group by tag_id
order by 2 desc
limit 5)

select tag_id,tag_name,tag_count from tags t inner join cte1 c on t.id=c.tag id



/*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

/*

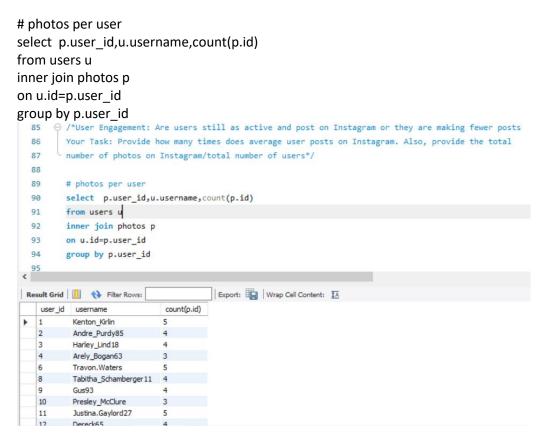
select dayname(created_at),count(*) from users group by dayname(created_at) order by 2 desc

#more generic approach

with cte1 as(select dayname(created_at) as day_name,count(*) as to from users group by dayname(created_at) order by 2 desc), cte2 as (select day_name,tc,dense_rank() over(order by tc desc) as 'rn1' from cte1) select * from cte2 where rn1=1

```
71
    L Your Task: What day of the week do most users register on? Provide insights on when to schedule an ad campaign */
72
73
     #more generic approach
74
75
   from users group by dayname(created_at) order by 2 desc),
77
    cte2 as (select day_name,tc,dense_rank() over(order by tc desc) as 'rn1' from cte1 )
     select * from cte2 where rn1=1
78
                         Export: Wrap Cell Content: IA
Result Grid | Filter Rows:
  day_name tc rn1
 Thursday 16
            1
 Sunday 16 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*/



#cte1 gives total photos,cte2gives total users,total query gives avg post per user
with cte1 as (SELECT count(id) as total_photos FROM photos),
cte2 as (SELECT count(id) as total_users FROM users)
select round(cte1.total_photos/cte2.total_users,2) as avg_post_per_user from cte1 inner join cte2

```
#cte1 gives total photos,cte2gives total users,total query gives avg post per user
with cte1 as (SELECT count(id) as total_photos FROM photos),

gete2 as (SELECT count(id) as total_users FROM users)

select round(cte1.total_photos/cte2.total_users,2) as avg_post_per_user from cte1 inner join cte2

100

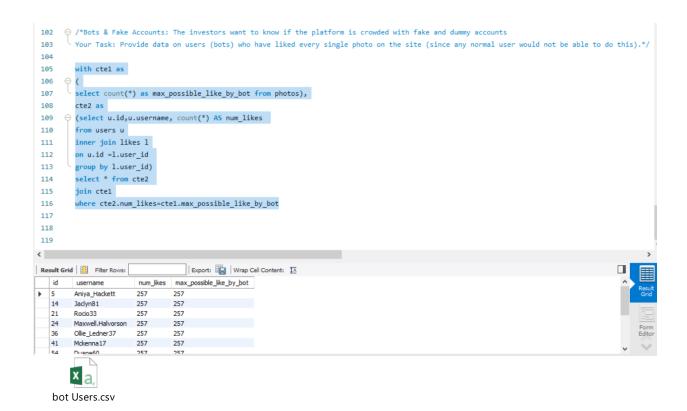
Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Wrap Cell Content: | Export: | 257
```



/*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).*/

```
with cte1 as
(
select count(*) as max_possible_like_by_bot from photos),
cte2 as
(select u.id,u.username, count(*) AS num_likes
from users u
inner join likes l
on u.id =l.user_id
group by l.user_id)
select * from cte2
join cte1
where cte2.num_likes=cte1.max_possible_like_by_bot
```



Result: I was able to create meaningful insights from the dataset which can be crucial for business. I was able to sharpen my SQL skills from here. Successfully completing this project boosted my confidence.