Instagram User Analytics

Project Description:

This project is an attempt to provide insights into questions asked by the management team by analysing the data provided by the product team of Instagram. To start with my analysis, I first understood the different metrics provided. I understood marketing and investor metrics. Then, I used the provided data for SQL analysis to find out the top-ranked oldest user, users who have never posted a single photo on Instagram, the most common hashtag, the total number of users, etc.

Approach:

Analysed the objective of every task and looked for the actual data that the team required. Then I imported the data into MySQL Workbench and ran numerous queries to provide insights for business benefits.

N.B.- As I was using MySQL for the first time, I made numerous mistakes. I could not understand the queries sometimes. I took the help of different resources. This was a challenging step. Finally, after Eshan's Session, I understood things more clearly.

Tech Stack Used:

MySQL Workbench 8.0

Result:

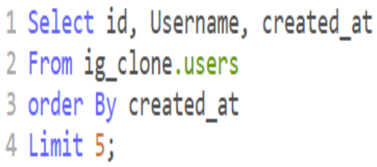
Marketing:

1. Rewarding Most Loyal Users:

People who have been using the platform for the longest time.

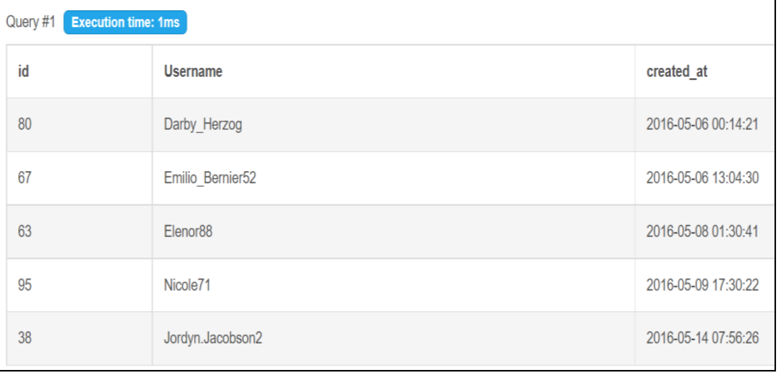
Task: Finding the 5 oldest users of Instagram from the database provided.

Query



```
Select id, Username, created_at
From ig_clone.users
order By created_at
Limit 5;
```

Output



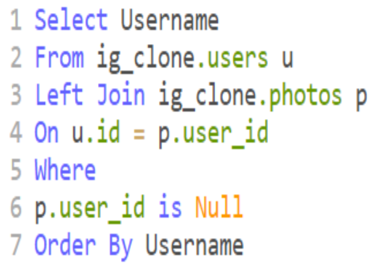
Query #1 Execution time: 1ms

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. Remind Inactive Users to Start Posting:

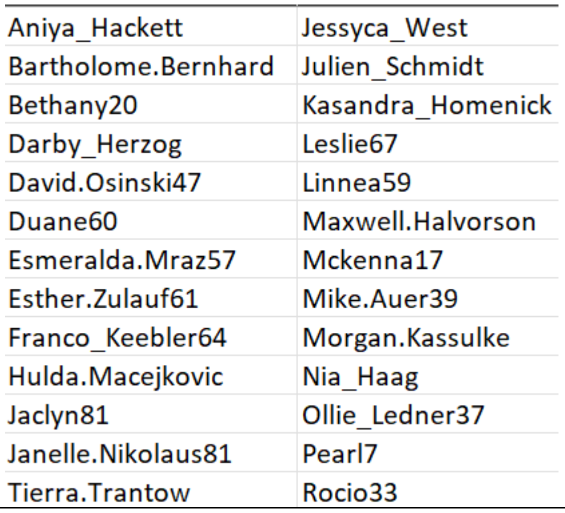
By sending them promotional emails to post their 1st photo.
Task: Finding the users who have never posted a single photo on Instagram

Query



```
Select Username
From ig_clone.users u
Left Join ig_clone.photos p
On u.id = p.user_id
Where
p.user_id is Null
Order By Username
```

Output

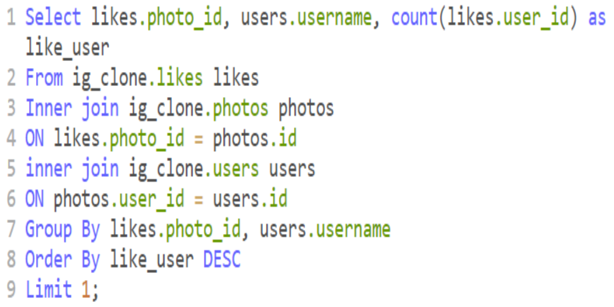


Aniya_Hackett	Jessyca_West
Bartholome.Bernhard	Julien_Schmidt
Bethany20	Kasandra_Homenick
Darby_Herzog	Leslie67
David.Osinski47	Linnea59
Duane60	Maxwell.Halvorson
Esmeralda.Mraz57	Mckenna17
Esther.Zulauf61	Mike.Auer39
Franco_Keebler64	Morgan.Kassulke
Hulda.Macejkovic	Nia_Haag
Jaclyn81	Ollie_Ledner37
Janelle.Nikolaus81	Pearl7
Tierra.Trantow	Rocio33

3. Declaring Contest Winner:

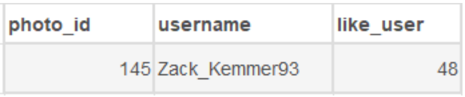
The team started a contest and the user who gets the most likes on a single photo will win the contest now they wish to declare the winner.
Task: Identifying the winner of the contest and provide their details to the team.

Query



```
Select likes.photo_id, users.username, count(likes.user_id) as
  like_user
From ig_clone.likes likes
Inner join ig_clone.photos photos
ON likes.photo_id = photos.id
inner join ig_clone.users users
ON photos.user_id = users.id
Group By likes.photo_id, users.username
Order By like_user DESC
Limit 1;
```

Output

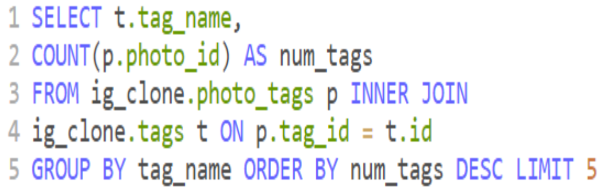


photo_id	username	like_user
145	Zack_Kemmer93	48

4. Hashtag Researching:

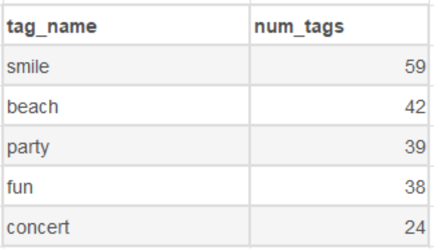
A partner brand wants to know, which hashtags to use in the post to reach the most people on the platform.
Task: Identifying and suggesting the top 5 most commonly used hashtags on the platform

Query



```sql
SELECT t.tag_name,
COUNT(p.photo_id) AS num_tags
FROM ig_clone.photo_tags p INNER JOIN
ig_clone.tags t ON p.tag_id = t.id
GROUP BY tag_name ORDER BY num_tags DESC LIMIT 5
```

Output

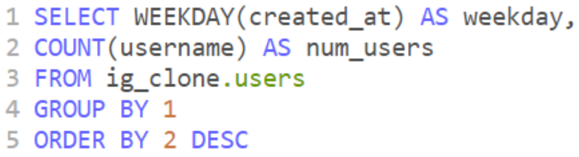


tag_name	num_tags
smile	59
beach	42
party	39
fun	38
concert	24

5. Launch AD Campaign:

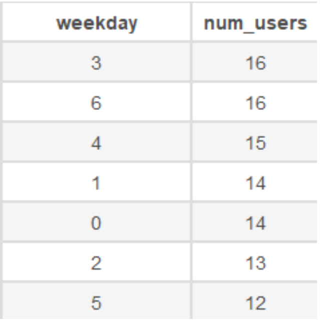
The team wants to know, which day would be the best day to launch ADs.
Task: What day of the week do most users register on? Provide insights on when to schedule an ad campaign

Query



```sql
SELECT WEEKDAY(created_at) AS weekday,
COUNT(username) AS num_users
FROM ig_clone.users
GROUP BY 1
ORDER BY 2 DESC
```

Output



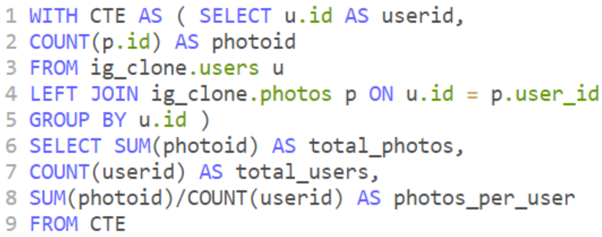
weekday	num_users
3	16
6	16
4	15
1	14
0	14
2	13
5	12

B) Investor Metrics:

1. User Engagement:

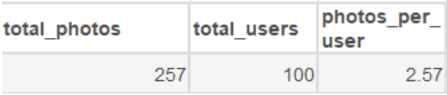
Are users still as active and post on Instagram or they are making fewer posts
Task: Provide how many times does average user posts on Instagram. Also, provide the total number of photos on Instagram/total number of users.

Query



```
WITH CTE AS ( SELECT u.id AS userid,
COUNT(p.id) AS photoid
FROM ig_clone.users u
LEFT JOIN ig_clone.photos p ON u.id = p.user_id
GROUP BY u.id )
SELECT SUM(photoid) AS total_photos,
COUNT(userid) AS total_users,
SUM(photoid)/COUNT(userid) AS photos_per_user
FROM CTE
```

Output

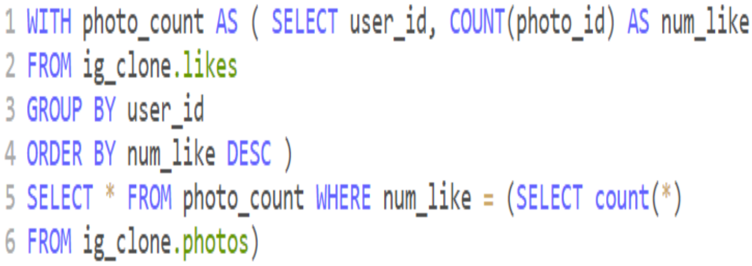


total_photos	total_users	photos_per_user
257	100	2.57

2. Bots & Fake Accounts:

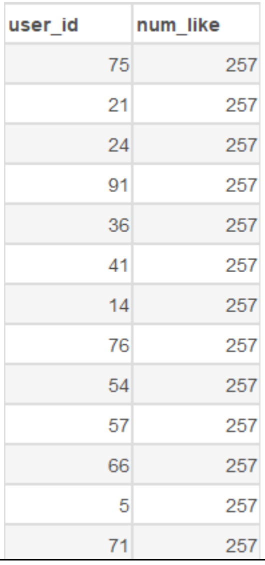
The investors want to know if the platform is crowded with fake and dummy accounts
Task: Provide data on users (bots) who have liked every single photo on the site (since any normal user would not be able to do this).

Query



```
WITH photo_count AS ( SELECT user_id, COUNT(photo_id) AS num_like
FROM ig_clone.likes
GROUP BY user_id
ORDER BY num_like DESC )
SELECT * FROM photo_count WHERE num_like = (SELECT count(*)
FROM ig_clone.photos)
```

Output



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