PROJECT - 2

Instagram User Analytics

SQL Fundamentals

A) Marketing Analysis:

1. **Loyal User Reward:** The marketing team wants to reward the most loyal users, i.e., those who have been using the platform for the longest time.

   Task: Identify the five oldest users on Instagram from the provided database.

   Query: select * from users order by created_at limit 5;

   Output:

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. **Inactive User Engagement:** The team wants to encourage inactive users to start posting by sending them promotional emails.

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

Query: select username from users
            left join photos
            on users.id = photos.user_id
            where photos.id is null;

Output:

username
Aniya_Hackett
Kasandra_Homenick
Jaclyn81
Rocio33
Maxwell.Halvorson
Tierra.Trantow
Pearl7
Ollie_Ledner37
Mckenna17
David.Osinski47
Morgan.Kassulke
Linnea59
Duane60
Julien_Schmidt
Mike.Auer39
Franco_Keebler64
Nia_Haag
Hulda.Macejkovic
Leslie67

Janelle.Nikolaus81
Darby_Herzog
Esther.Zulauf61
Bartholome.Bernhard
Jessyca_West
Esmeralda.Mraz57
Bethany20

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

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

Query: select username,photos.id,photos.image_url,
count(likes.user_id) as total from photos
inner join likes on likes.photo_id = photos.id
inner join users on photos.user_id = users.id
group by photos.id order by total desc
limit 1;

Output:

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

4. Hashtag Research: A partner brand wants to know the most popular hashtags to use in their posts to reach the most people.

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

Query: select tags.tag_name,count(*)as total_tags
from photo_tags join tags
on photo_tags.tag_id = tags.id
group by tags.id
order by total_tags desc
limit 5;

Output:

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

5. **Ad Campaign Launch:** The team wants to know the best day of the week to launch ads.

Task: Determine the day of the week when most users register on Instagram. Provide insights on when to schedule an ad campaign.

Query: select dayname(created_at) as day,count(*)
as total from users group by day
order by total desc limit 1;

Output:

day	total
Thursday	16

B) Investor Metrics:

1. **User Engagement:** Investors want to know if users are still active and posting on Instagram or if they are making fewer posts.

   Task: 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.

   Query: select (select count(*) from photos) / (select count(*) from users) as avg;

   Output:

   avg
   2.5700

2. **Bots & Fake Accounts:** Investors want to know if the platform is crowded with fake and dummy accounts.

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

   Query: select user_id,count(*) as num_likes from likes
   group by user_id
   having num_likes = (select count(*) from photos);
   select u.username,count(*)as num_likes
   from users u join likes i on u.id = i.user_id
   group by u.id
   having num_likes = (select count(*) from photos);

Output:

username	num_likes
Aniya_Hackett	257
Jaclyn81	257
Rocio33	257
Maxwell.Halvorson	257
Ollie_Ledner37	257
Mckenna17	257
Duane60	257
Julien_Schmidt	257
Mike.Auer39	257
Nia_Haag	257
Leslie67	257
Janelle.Nikolaus81	257
Bethany20	257