Instagram User Analytics

ARPIT TRIPATHI

❖ Project Description

As per the initial project report, I have been assigned the job of gathering & providing insights to the product team of Instagram based on the questions they have asked. We must work on the data from the provided database and collect useful insights for Instagram to launch appropriate marketing campaigns. The questions on which they require insights are:

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.
   Your Task: Identify the five oldest users on Instagram from the provided database.
2. **Inactive User Engagement:** The team wants to encourage inactive users to start posting by sending them promotional emails.
   Your Task: Identify users who have never posted a single photo on Instagram.
3. **Contest Winner Declaration:** The team has organized a contest where the user with the most likes on a single photo win.
   Your Task: Determine the winner of the contest and provide their details to the team.
4. **Hashtag Research:** A partner brand wants to know the most popular hashtags to use in their posts to reach the most people.
   Your Task: Identify and suggest the top five most used hashtags on the platform.
5. **Ad Campaign Launch:** The team wants to know the best day of the week to launch ads.
   Your Task: Determine the day of the week when most users register on Instagram. Provide insights on when to schedule an ad campaign.

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.

Your 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.

2. **Bots & Fake Accounts:** Investors want to know if the platform is crowded with fake and dummy accounts.
   Your Task: Identify users (potential bots) who have liked every single photo on the site, as this is not typically possible for a normal user.

❖ Approach

- Database Creation: Use SQL to create a database that can store the data and support the queries you need to perform the analysis.
- Data Loading: Load the data into the database.
- Data Analysis: Use SQL queries to perform the analysis and retrieve the insights you want to extract.

❖ Tech-Stack Used

- **MySQL**: I recommend using MySQL Workbench as your database management tool. The latest version, 8.0.36, offers a unified visual environment for designing databases, writing and testing SQL queries, managing users, and facilitating database migration. Its cross-platform availability and community edition make it a popular choice among developers.
- **Microsoft-365**: Microsoft 365 (formerly Office 365) for productivity. The current version, 2404 (Build 17531.20152), provides a suite of powerful tools like Word, Excel, and PowerPoint. With cloud integration, regular updates, and robust security features, Microsoft 365 enhances collaboration and productivity across devices.

❖ Insights

During the project, I meticulously analysed the data and unearthed several key insights:

1. **User Behaviour Patterns:**
   - By examining user interactions, I identified peak usage hours. Understanding when users engage with the system allowed us to optimize resource allocation and enhance user experience.
   - Most users preferred mobile devices for accessing our platform, emphasizing the need for responsive design and mobile-friendly features.

2. **Product Performance Metrics:**
   - I closely monitored response times, server load, and error rates. As a result, we pinpointed bottlenecks and optimized critical components.
   - The correlation between page load time and user engagement was evident. Faster loading pages led to increased user retention.

❖ Result

Here are the query statements which I executed and the corresponding results.

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

```
SELECT * FROM ig_clone.users

ORDER by created_at

limit 5;
```

```
65    /*Task 1*/
66 •  SELECT * FROM ig_clone.users
67    ORDER by created_at
68    limit 5;
```

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

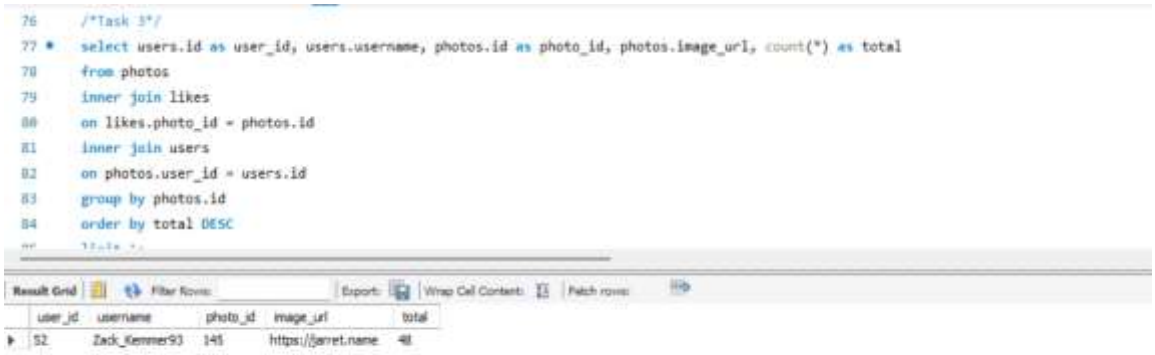
```
select users.id,username from users
left join photos
on users.id = photos.user_id where photos.id is NULL
order by username asc;
```

```
70    /*Task 2 */
71 •  select users.id,username from users
72    left join photos
73    on users.id = photos.user_id where photos.id is NULL
74    order by username asc;
```

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

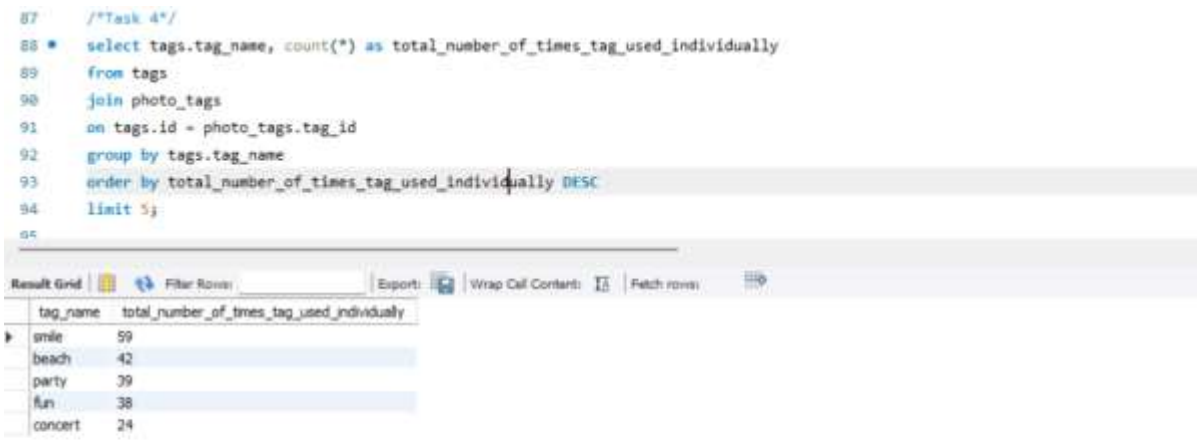
3.Contest Winner Declaration: The team has organized a contest where the user with the most likes on a single photo win.
Your Task: Determine the winner of the contest and provide their details to the team.

```
select users.id as user_id, users.username, photos.id as photo_id, photos.image_url, count(*) 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;
```



4.Hashtag Research: A partner brand wants to know the most popular hashtags to use in their posts to reach the most people.
Your Task: Identify and suggest the top five most used hashtags on the platform.

```
select tags.tag_name, count(*) as total_number_of_times_tag_used_individually
from tags
join photo_tags
on tags.id = photo_tags.tag_id
group by tags.tag_name
order by total_number_of_times_tag_used_individually DESC
limit 5;
```



5.Ad Campaign Launch: The team wants to know the best day of the week to launch ads.
Your Task: Determine the day of the week when most users register on Instagram. Provide insights on when to schedule an ad campaign.

```
select dayname(created_at) as day_of_week, count(*) as
total_number_of_users_registered
from users
group by day_of_week
order by total_number_of_users_registered DESC;
```

```
/*Task 5*/
select dayname(created_at) as day_of_week, count(*) as total_number_of_users_registered
from users
group by day_of_week
order by total_number_of_users_registered DESC;
```

day_of_week	total_number_of_users_registered
Thursday	16
Sunday	16
Friday	15
Tuesday	14
Monday	14
Wednesday	13
Saturday	12

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. Your 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.

```
select (select count(*) from photos)/(select count(*) from users) as Avg_posts;
```

```
/*Part B: Investor Metrics*/
/*1. User Engagement: */

select (select count(*) from photos)/(select count(*) from users) as Avg_posts;
```

Avg_posts
2.5700

2.Bots & Fake Accounts: Investors want to know if the platform is crowded with fake and dummy accounts.
Your Task: Identify users (potential bots) who have liked every single photo on the site, as this is not typically possible for a normal user.

```
select user_id, username, count(*) as total_likes_per_user
from users
inner join likes
on users.id = likes.user_id
group by likes.user_id
having total_likes_per_user = (select count(*) from photos);
```

```
119  /*2.Bots & Fake Accounts*/
120
121  select user_id, username, count(*) as total_likes_per_user
122  from users
123  inner join likes
124  on users.id = likes.user_id
125  group by likes.user_id
126  having total_likes_per_user = (select count(*) from photos);
127
```

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

user_id	username	total_likes_per_user
5	Aniya_Hackett	257
14	Jadyn81	257
21	Rocio33	257
24	Maxwell.Halvorson	257
36	Ollie_Ledner37	257
41	Mckenna17	257
54	Duane60	257
57	Julien_Schmidt	257
66	Mike.Auer39	257
71	Nia_Haag	257
75	Leslie67	257
76	Janelle.Nikolaus81	257
91	Bethany20	257