Project on
Instagram User Analytics

Presented by:

Bhawana yadav

Project Description

- In this project I am analyzing user interactions and engagement with the Instagram app to provide valuable insights that can help the business grow.
- User analysis involves tracking how users engage with a digital product, such as a software application or a mobile app. The insights derived from this analysis can be used by various teams within the business. For example, the marketing team might use these insights to launch a new campaign.
- In this project, I am using MySQL Workbench as tool to analyze Instagram user data and answer questions posed by the management team. My insights will help the product manager and the rest of the team make informed decisions about the future direction of the Instagram app.
- the goal of this project is to use SQL skills to extract meaningful insights from the data. My findings could potentially influence the future development of one of the world's most popular social media platforms.

Approach

Approach / Steps that are taken by me to analyze data and execute question:

1. First of all I tried to understand the project and each question of it. Then I solved them serial wise with help of SQL commands.
2. There are 2 tasks into the project :
    1. **Marketing Analysis:**
        1. **Task:** Identify the five oldest users on Instagram from the provided database.

Solution:

Query:

```
SELECT
    username
FROM
    users
ORDER BY created_at
LIMIT 5;
```

Output:

	username
▶	Darby_Herzog
	Emilio_Bernier52
	Elenor88
	Nicole71
	Jordyn.Jacobson2

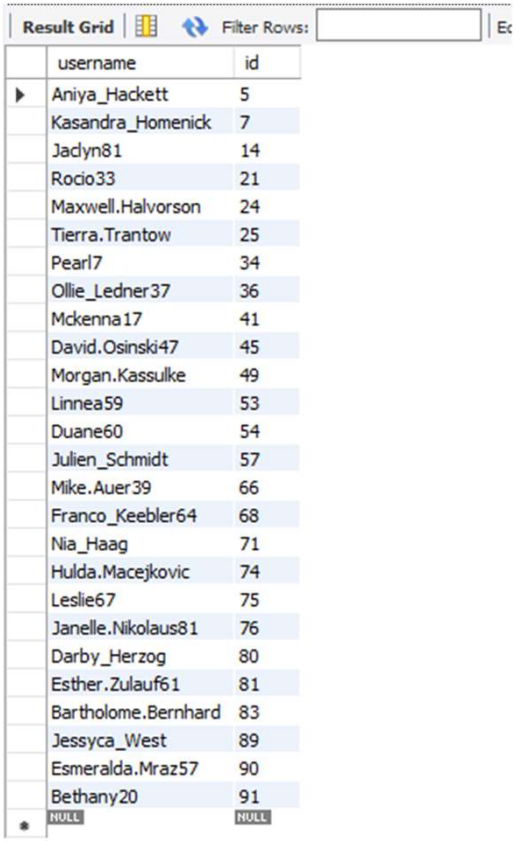
2. **Task:** Identify users who have never posted a single photo on Instagram.

Solution:

Query:

```
SELECT
    username, id
FROM
    users
WHERE
    id NOT IN (SELECT
            user_id
        FROM
            photos
        WHERE
            users.id = photos.user_id);
```

Output:



Result Grid | Filter Rows:

username	id
Aniya_Hackett	5
Kasandra_Homenick	7
Jadyn81	14
Rocio33	21
Maxwell.Halvorson	24
Tierra.Trantow	25
Pearl7	34
Ollie_Ledner37	36
Mckenna17	41
David.Osinski47	45
Morgan.Kassulke	49
Linnea59	53
Duane60	54
Julien_Schmidt	57
Mike.Auer39	66
Franco_Keebler64	68
Nia_Haag	71
Hulda.Macejkovic	74
Leslie67	75
Janelle.Nikolaus81	76
Darby_Herzog	80
Esther.Zulauf61	81
Bartholome.Bernhard	83
Jessyca_West	89
Esmeralda.Mraz57	90
Bethany20	91
NULL	NULL

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

Solution:

Query:

```
WITH PhotoLikeCounts AS (
    SELECT p.id, p.user_id, COUNT(l.photo_id) AS like_count
    FROM photos as p
    LEFT JOIN likes as l ON p.id = l.photo_id
    GROUP BY p.id, p.user_id
),
WinningPhoto AS (
    SELECT id, user_id, like_count
    FROM PhotoLikeCounts
    ORDER BY like_count desc
    LIMIT 1
)
SELECT u.id, u.username, wp.like_count
FROM users as u
JOIN WinningPhoto as wp ON u.id = wp.user_id;
```

Output:

Result Grid | Filter Rows: | Export: | Wr

id	username	like_count
52	Zack_Kemmer93	48

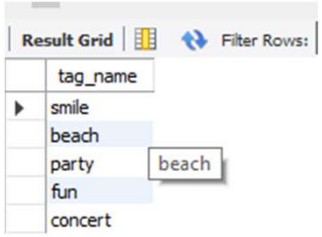
4. **Task:** Identify and suggest the top five most commonly used hashtags on the platform.

Solution:

Query:

```
SELECT
    tags.tag_name
FROM
    (SELECT
        tag_id, COUNT(tag_id) AS more_used
    FROM
        photo_tags
    GROUP BY tag_id
    ORDER BY more_used DESC
    LIMIT 5) new_tab
        inner JOIN
    tags ON new_tab.tag_id = tags.id;
```

Output:



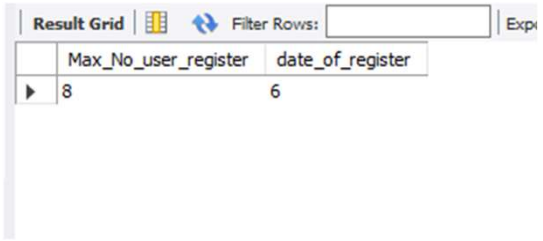
5. **Task:** Determine the day of the week when most users register on Instagram. Provide insights on when to schedule an ad campaign.

Solution:

Query:

```
SELECT
    COUNT(*) AS Max_No_user_register,
    EXTRACT(DAY FROM created_at) AS date_of_register
FROM
    users
GROUP BY date_of_register
ORDER BY Max_No_user_register DESC
LIMIT 1;
```

Output:



2. Investor Metrics:

1. **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.

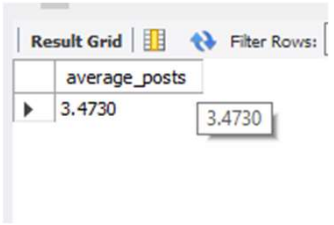
Solution:

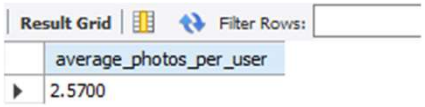
Query:

```
SELECT
    AVG(count) AS average_posts
FROM
    (SELECT
        user_id, COUNT(id) AS count
    FROM
        photos
    GROUP BY user_id) AS new;
```

```
SELECT
    (SELECT
        COUNT(id)
    FROM
        photos) / (SELECT
        COUNT(id)
    FROM
        users) AS average_photos_per_user;
```

Output:





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

Solution:

Query:

```
WITH TotalPhotos AS (
    SELECT COUNT(*) AS photo_count
    FROM photos
),
UserLikeCounts AS (
    SELECT user_id, COUNT(*) AS like_count
    FROM likes
    GROUP BY user_id
)
SELECT u.id, u.username
FROM users as u
JOIN UserLikeCounts as ulc ON u.id = ulc.user_id
JOIN TotalPhotos as tp ON ulc.like_count = tp.photo_count
WHERE ulc.like_count = tp.photo_count;
```

Output:

Result Grid | Filter Rows: | Export:

id	username
5	Aniya_Hackett
14	Jaclyn81
21	Rocio33
24	Maxwell.Halvorson
36	Ollie_Ledner37
41	Mckenna17
54	Duane60
57	Julien_Schmidt
66	Mike.Auer39
71	Nia_Haag
75	Leslie67
76	Janelle.Nikolaus81
91	Bethany20

Tech-Stack Used

- I am using MySQL Workbench for Project
- **Version of software:**

- I can view the account information of all users on the MySQL server.
- MySQL Workbench gives access to add and remove users.
- MySQL Workbench grants and revokes privileges.
- I can modify global and database permissions on the MySQL server.
- I can change passwords using MySQL.
- It is easy to use. I can easily store , retrieve , modify and delete data from database easily.

Insights

Insights on the questions of projects:

A) Marketing Analysis:

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

Insights: Out of 100 users there are only I have found 5 oldest users.

But out of these 100 the most oldest user ' **Darby_Herzog'**

is never posted a single photo on Instagram.

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

Insights: The most popular hashtags are smile ,party , concert etc. These are more helpful for fashion brand like they can show these hashtag or images on t-shirts , pants etc.

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

Insights: Out of 100 users only 8 users register on same day (6 th)

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

Insights: According to average 3.437 , these are user who are active and using the Instagram .With the help of these users the brands can do promotion , etc.

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

Insights: Out of 100 there are 13 users who are fake and have dummy accounts.

And all of these never posted a single post on Instagram but they liked every post .

Results

- With the help of all the Insights , it is conclude that some users are fake on Instagram and also there is one user who is oldest on app but he never posted a single post.
- This project helps me to understand task /query questions from the point of view of real world tasks. I have learnt nested queries , joins, Date function and performed my SQL skills on actual data.