**SOCIAL MEDIA ANALYSIS**

**You are hired as a data analyst at Meta and asked to collaborate with the Marketing team. Marketing teams want to leverage Instagram's user data to develop targeted marketing strategies that will increase user engagement, retention, and acquisition. Provide insights and recommendations to address the following objectives.**

Objective Questions

1. Are there any tables with duplicate or missing null values? If so, how would you handle them?

SELECT \* FROM tags

WHERE id IS NULL OR tag\_name IS NULL OR created\_at IS NULL;

SELECT \* FROM users

WHERE id IS NULL OR username IS NULL OR created\_at IS NULL;

SELECT \* FROM photos

WHERE id IS NULL OR image\_url IS NULL OR user\_id IS NULL OR created\_dat IS NULL;

SELECT \* FROM photo\_tags

WHERE photo\_id IS NULL OR tag\_id IS NULL;

SELECT \* FROM likes

WHERE user\_id IS NULL OR photo\_id IS NULL OR created\_at IS NULL;

SELECT \* FROM follows

WHERE follower\_id IS NULL OR followee\_id IS NULL OR created\_at IS NULL;

SELECT \* FROM comments

WHERE id IS NULL OR comment\_text IS NULL OR user\_id IS NULL OR photo\_id IS NULL OR created\_at IS NULL;

1. What is the distribution of user activity levels (e.g., number of posts, likes, comments) across the user base?

With CommentsCTE as

(

select u.id as UserID,count(c.id) as TotalComments

from comments c

join users u on u.id=c.user\_id

group by u.id

order by u.id

),

LikesCTE as

(

select u.id as UserID,count(\*) as Totallikes

from likes l

join users u on u.id=l.user\_id

group by u.id

order by u.id

),

PostCTE as

(

select u.id as UserID,count(\*) as TotalPosts

from photos p

join users u on u.id=p.user\_id

group by u.id

order by u.id

)

select

u.id as UserID,

TotalComments,

TotalLikes,

TotalPosts

from users u

join CommentsCTE c on u.id=c.UserID

join LikesCTE lc on u.id=lc.UserID

join PostCTE pc on u.id=pc.UserID;

* First create 3 different CTE to count Comments,Likes and Posts for each user.
* In Each CTE use that table for comment,likes and photos and did join them with users table on userId column.
* Lastly,joined all CTE with users table on userId column to get desired output as single table.

1. Calculate the average number of tags per post (photo\_tags and photos tables).

SELECT ROUND(AVG(tag\_count),2) AS avg\_tags\_per\_post

FROM

(SELECT photo\_id,COUNT(DISTINCT tag\_id) AS tag\_count

FROM photo\_tags

GROUP BY 1

) count\_tags\_per\_post;

* The **subquery** groups all tags by photo\_id and counts the distinct tags for each photo.
* The **main query** calculates the average of these tag counts and rounds the result to two decimal places.

1. Identify the top users with the highest engagement rates (likes, comments) on their posts and rank them.

WITH likes\_comments AS (

SELECT p.id, p.user\_id,

COUNT(DISTINCT c.id) AS count\_comments,

COUNT(DISTINCT l.user\_id) AS count\_likes

FROM photos p

LEFT JOIN comments c

ON p.id = c.photo\_id

LEFT JOIN likes l

ON p.id = l.photo\_id

GROUP BY p.id

),

user\_likes\_comments AS (

SELECT user\_id,

SUM(count\_likes) AS total\_likes,

SUM(count\_comments) AS total\_comments,

SUM(count\_comments + count\_likes) AS total\_engagement

FROM likes\_comments

GROUP BY user\_id

)

SELECT u.user\_id, u1.username, u.total\_likes, u.total\_comments, u.total\_engagement,

RANK() OVER(ORDER BY total\_engagement DESC) AS user\_rank

FROM user\_likes\_comments u

JOIN users u1

ON u.user\_id = u1.id

ORDER BY total\_engagement DESC;

* Created first CTE for photos to check how many likes and comments are there on that photo. I used photos,comments and likes table and joined them on photo\_id to get count of comments and likes per photo.
* For 2nd CTE, I used first CTE to get total likes and comments for that user. And did sum of likes and comments to get total engagement for that particular user.
* At last in select query rank all the user on their enganement level from higher to lower and display them from highest to lowest.

1. Which users have the highest number of followers and followings?

WITH number\_of\_followers AS (

SELECT follower\_id, COUNT(follower\_id) AS follower\_count

FROM follows

GROUP BY follower\_id

),

number\_of\_followings AS (

SELECT followee\_id, COUNT(followee\_id) AS followee\_count

FROM follows

GROUP BY followee\_id

)

SELECT u.id, u.username,

MAX(follower\_count) AS max\_followers,

MAX(followee\_count) AS max\_followings

FROM users u

LEFT JOIN number\_of\_followers fr

ON u.id = fr.follower\_id

LEFT JOIN number\_of\_followings fe

ON u.id = fe.followee\_id

GROUP BY u.id, u.username

ORDER BY max\_followers DESC, max\_followings DESC, u.id;

* In the above query we are taking two cte first is number\_of\_followers and number\_of\_followings ,In number\_of\_followers are counting follower\_id as number of follower after that group by follower\_id ,
* In number\_of\_followings cte ,we are counting followee \_id AS followee\_count and group by followee\_id ,after that using both cte finding max of max\_followers,and max\_followings with join both cte .

1. Calculate the average engagement rate (likes, comments) per post for each user.

WITH cte AS

(

SELECT

p.id AS photo\_id,

p.user\_id,

count(DISTINCT l.user\_id) AS like\_count,

count(DISTINCT c.id) AS comment\_count,

count(DISTINCT l.user\_id) + COUNT(DISTINCT c.id) AS total\_engagement

FROM photos p

LEFT JOIN likes l on p.id = l.photo\_id

LEFT JOIN comments c on p.id = c.photo\_id

GROUP BY p.id

),

cte2 AS

(

SELECT

pe.user\_id,

sum(pe.total\_engagement) AS total\_engagement,

count(pe.photo\_id) AS post\_count

FROM cte pe

GROUP BY pe.user\_id

)

SELECT

ue.user\_id,

u.username,

ue.total\_engagement,

ue.post\_count,

round((ue.total\_engagement / ue.post\_count),2) AS average\_engagement\_per\_post

FROM cte2 ue

JOIN users u ON ue.user\_id = u.id

ORDER BY ue.user\_id;

* Created first CTE for photos to check how many likes and comments are there on that photo.I used photos,comments and likes table and joined them on photo\_id to get count of comments and likes per photo.
* For 2nd CTE,I used first CTE to get total likes and comments for that user.And did sum of likes and comments to get total engagement for that particular photo.
* At last in select query calculated average engagement of likes and comments per photo.

1. Get the list of users who have never liked any post (users and likes tables)

SELECT id,username

FROM users WHERE id NOT IN (SELECT user\_id FROM likes);

* From users table selected id, username, and put condition of returning only those id which are not in likes table.

1. How can you leverage user-generated content (posts, hashtags, photo tags) to create more personalized and engaging ad campaigns?

WITH cte AS(

SELECT u.id AS user\_id,

username,

t.id AS tag\_id,

tag\_name,

COUNT(DISTINCT l.user\_id) AS likes,

COUNT(DISTINCT c.user\_id) AS comments,

DENSE\_RANK() OVER(PARTITION BY u.id ORDER BY (COUNT(DISTINCT l.user\_id)+COUNT(DISTINCT c.user\_id)) DESC) AS engagement\_rank

FROM users u

JOIN photos p ON u.id=p.user\_id

JOIN photo\_tags pt ON p.id=pt.photo\_id

JOIN tags t ON t.id=pt.tag\_id

JOIN likes l ON pt.photo\_id=l.photo\_id

JOIN comments c ON pt.photo\_id=c.photo\_id

GROUP BY 1,2,3,4

)

SELECT user\_id, username, tag\_name FROM cte

WHERE engagement\_rank=1;

* We have ranked by likes and comments on different tags by different to see what are people most engaging in.
* We can conclude the specific tags most liked my different people, to increase the user engagement by showing more of those posts, reels.

1. Are there any correlations between user activity levels and specific content types (e.g., photos, videos, reels)? How can this information guide content creation and curation strategies?

SELECT t.id AS tag\_id, tag\_name,

COUNT(DISTINCT pt.photo\_id) AS photo\_id,

COUNT(DISTINCT l.user\_id) AS likes,

COUNT(DISTINCT c.user\_id) AS comments,

DENSE\_RANK() OVER(ORDER BY COUNT(DISTINCT pt.photo\_id) DESC,COUNT(DISTINCT l.user\_id)DESC,COUNT(DISTINCT c.user\_id)) AS engagement\_rank

FROM tags t

LEFT JOIN photo\_tags pt ON t.id=pt.tag\_id

LEFT JOIN likes l ON pt.photo\_id=l.photo\_id

LEFT JOIN comments c ON pt.photo\_id=c.photo\_id

GROUP BY 1,2;

* Yes, there are often correlations between user activity levels and specific content types. For example, some users may prefer photos over videos, while others may engage more with reels or short-form video content. By analyzing user activity data, content creators and curators can identify which types of content are resonating most with their audience and tailor their content creation and curation strategies accordingly.
* In the query we can see the tags that user mostly engage with.

1. Calculate the total number of likes, comments, and photo tags for each user.

WITH CommentCTE AS

(

SELECT

u.id,

COUNt(\*) AS totalcomments

FROM users u

JOIN comments c ON c.user\_id=u.id

GROUP BY u.id

ORDER BY u.id DESC

),

LikesCTE AS

(

SELECT

u.id,

COUNt(\*) AS totallikes

FROM users u

JOIN likes l ON l.user\_id=u.id

GROUP BY u.id

ORDER BY u.id DESC

),

PhototagCTE AS

(

SELECT

p.user\_id,

COUNT(\*) AS totalPhototags

from photos p

join photo\_tags p1 ON p1.photo\_id=p.id

GROUP BY p.user\_id

ORDER BY p.user\_id DESC

)

SELECT

user\_id,

totalcomments,

totallikes,

totalPhototags

FROM CommentCTE c

JOIN LikesCTE l ON l.id=c.id

JOIN PhototagCTE p ON p.user\_id=c.id;

* First create 3 different CTE to count Comments,Likes and phototags for each user.
* In Each CTE use that table for comment,likes and photottags and did join them with users table on userId column.
* Lastly,joined all CTE on userId column to get desired output as single table.

1. Rank users based on their total engagement (likes, comments, shares) over a month.

WITH LikesCount AS

(

SELECT p.user\_id,

COUNT(l.photo\_id) AS total\_likes

FROM photos p

LEFT JOIN likes l ON p.id = l.photo\_id

WHERE l.created\_at >= NOW() - INTERVAL 1 MONTH

GROUP BY p.user\_id

),

CommentsCount AS

(

SELECT

p.user\_id,

COUNT(c.photo\_id) AS total\_comments

FROM photos p

LEFT JOIN comments c ON p.id = c.photo\_id

WHERE c.created\_at >= NOW() - INTERVAL 1 MONTH

GROUP BY p.user\_id

),

Engagement AS

(

SELECT

u.id AS user\_id,

u.username,

COALESCE(l.total\_likes, 0) + COALESCE(c.total\_comments, 0) AS total\_engagement

FROM users u

LEFT JOIN LikesCount l ON u.id = l.user\_id

LEFT JOIN CommentsCount c ON u.id = c.user\_id

)

SELECT

user\_id,

username,

total\_engagement,

RANK() OVER (ORDER BY total\_engagement DESC) AS engagement\_rank

FROM Engagement

ORDER BY engagement\_rank;

* Created 2 different CTE to count no. of likes and comments in last month. 3rd CTE for to count total engagement of user. In select query,used rank function to display their rank from highest to lowest on their engagement level.

1. Retrieve the hashtags that have been used in posts with the highest average number of likes. Use a CTE to calculate the average likes for each hashtag first.

WITH tag\_name\_and\_likes AS (

SELECT t.id tag\_id,

tag\_name,

pt.photo\_id,

COUNT( DISTINCT l.user\_id) total\_likes,

AVG(COUNT( DISTINCT l.user\_id)) OVER(PARTITION BY t.id) AS avg\_likes

FROM tags t

LEFT JOIN photo\_tags pt ON t.id=pt.tag\_id

JOIN likes l ON l.photo\_id=pt.photo\_id

GROUP BY 1,2,3

)

SELECT DISTINCT tag\_id,tag\_name

FROM tag\_name\_and\_likes

WHERE avg\_likes IN (SELECT MAX(avg\_likes) FROM tag\_name\_and\_likes)

ORDER BY 1;

* In above CTE we have calculated avg likes per tags.
* In the main query we selected the tag that has maximum average likes.
* We can conclude that tag named dreamy has most used in posts.

1. Retrieve the users who have started following someone after being followed by that person

SELECT f1.follower\_id AS user1\_as\_follower,

f1.followee\_id AS user2\_as\_following,

f1.created\_at as followed\_at,

f2.follower\_id AS user2\_as\_follower,

f2.followee\_id AS user1\_as\_following,

f2.created\_at AS followed\_back\_at

FROM follows f1

JOIN follows f2 ON f1.followee\_id=f2.follower\_id

AND f1.follower\_id=f2.followee\_id

WHERE f2.created\_at<f1.created\_at

ORDER BY 1;

* Using this query no data is found as created date in follows table is same for all.

Subjective Questions

1. Based on user engagement and activity levels, which users would you consider the most loyal or valuable? How would you reward or incentivize these users?

WITH cte1 AS

(

SELECT user\_id, COUNT(\*) AS likes\_given

FROM likes

WHERE created\_at >= NOW() - INTERVAL 1 MONTH

GROUP BY user\_id

),

cte2 AS

(

SELECT p.user\_id, COUNT(\*) AS likes\_received

FROM likes l

JOIN photos p ON l.photo\_id = p.id

WHERE l.created\_at >= NOW() - INTERVAL 1 MONTH

GROUP BY p.user\_id

),

cte3 AS

(

SELECT user\_id, COUNT(\*) AS comments\_given

FROM comments

WHERE created\_at >= NOW() - INTERVAL 1 MONTH

GROUP BY user\_id

),

cte4 AS

(

SELECT p.user\_id, COUNT(\*) AS comments\_received

FROM comments c

JOIN photos p ON c.photo\_id = p.id

WHERE c.created\_at >= NOW() - INTERVAL 1 MONTH

GROUP BY p.user\_id

),

cte5 AS

(

SELECT user\_id, COUNT(\*) AS photos\_uploaded

FROM photos

WHERE created\_dat >= NOW() - INTERVAL 1 MONTH

GROUP BY user\_id

),

cte6 AS

(

SELECT followee\_id AS user\_id, COUNT(\*) AS followers

FROM follows

GROUP BY followee\_id

),

cte7 AS

(

SELECT follower\_id AS user\_id, COUNT(\*) AS followings

FROM follows

GROUP BY follower\_id

),

Engagement AS

(

SELECT

u.id AS user\_id,

u.username,

COALESCE(lg.likes\_given, 0) + COALESCE(lr.likes\_received, 0) +

COALESCE(cg.comments\_given, 0) + COALESCE(cr.comments\_received, 0) +

COALESCE(pu.photos\_uploaded, 0) + COALESCE(f.followers, 0) +

COALESCE(fg.followings, 0) AS total\_engagement

FROM

users u

LEFT JOIN cte1 lg ON u.id = lg.user\_id

LEFT JOIN cte2 lr ON u.id = lr.user\_id

LEFT JOIN cte3 cg ON u.id = cg.user\_id

LEFT JOIN cte4 cr ON u.id = cr.user\_id

LEFT JOIN cte5 pu ON u.id = pu.user\_id

LEFT JOIN cte6 f ON u.id = f.user\_id

LEFT JOIN cte7 fg ON u.id = fg.user\_id

)

SELECT

user\_id,

username,

total\_engagement,

RANK() OVER (ORDER BY total\_engagement DESC) AS engagement\_rank

FROM Engagement

ORDER BY engagement\_rank;

* Loyal and high-value users are those who engage consistently with the platform, contribute valuable content, participate in discussions, and have a positive influence on the community. These users are often the ones who offer feedback, assist other users, and show a strong, ongoing commitment to the platform.
* To keep these loyal users motivated, you might want to introduce a loyalty program with exclusive benefits such as access to premium features, discounts, early access to updates, or public recognition through community titles, badges, or profile highlights. Physical rewards like branded merchandise or gift cards could also be considered, alongside organizing special events or challenges that create a sense of importance for these top users.
* Ultimately, the idea is to acknowledge their efforts and offer rewards that are meaningful to them. This not only shows appreciation for their involvement but also keeps them engaged and encourages continued participation. By rewarding their loyalty, you create a deeper connection that encourages them to stay active and invested in the platform for the long term.

1. For inactive users, what strategies would you recommend to re-engage them and encourage them to start posting or engaging again?

* All users are active users.

WITH LastActivity AS

(

SELECT

u.id AS user\_id,

MAX(GREATEST(

(l.created\_at),

(c.created\_at),

(p.created\_dat)

)) AS last\_activity

FROM users u

LEFT JOIN likes l ON u.id = l.user\_id

LEFT JOIN comments c ON u.id = c.user\_id

LEFT JOIN photos p ON u.id = p.user\_id

GROUP BY u.id

)

SELECT

u.id,

u.username,

la.last\_activity

FROM users u

JOIN LastActivity la ON u.id = la.user\_id

WHERE la.last\_activity < NOW() - INTERVAL 3 MONTH;

**Recommendations:**

* **Personalized outreach**: Reach out to inactive users with personalized messages, thanking them for their past contributions and inviting them back to the community. Ask for feedback on why they’ve been less active and how you can improve their experience to encourage re-engagement.
* **Exclusive incentives**: Provide special deals or rewards for users who haven’t engaged recently, such as discounts, free trials, or access to premium features. These exclusive offers can motivate them to start contributing again.
* **Create urgency**: Introduce limited-time promotions or time-sensitive deals that motivate inactive users to return and engage quickly, making them feel they’ll miss out if they delay.
* **Showcase new updates**: Highlight new features or content that have been added to the platform since their last visit. This can spark curiosity and encourage users to return and explore what’s new.
* **Host contests or challenges**: Launch interactive contests or challenges that require participation, offering a fun way for inactive users to re-engage with the platform and the community.
* **Deliver engaging content**: Ensure the platform offers fresh, relevant, and engaging content that resonates with your audience. High-quality content is key to bringing inactive users back and keeping them involved.
* **Offer personalized recommendations**: Use data to provide personalized content recommendations to inactive users, showcasing posts, discussions, or features that align with their previous behavior and interests, making the platform feel more tailored to them.
* **Reconnect through surveys or polls**: Send out brief surveys or polls to inactive users asking for their input on what could make the platform more valuable to them. Not only does this invite feedback, but it also re-establishes communication and signals that their opinion is valued.

1. Which hashtags or content topics have the highest engagement rates? How can this information guide content strategy and ad campaigns?

WITH PhotoEngagement AS

(

SELECT

p.id AS photo\_id,

COALESCE(COUNT(DISTINCT l.user\_id), 0) + COALESCE(COUNT(DISTINCT c.id), 0) AS total\_engagement

FROM photos p

LEFT JOIN likes l ON p.id = l.photo\_id

LEFT JOIN comments c ON p.id = c.photo\_id

GROUP BY p.id

),

HashtagEngagement AS

(

SELECT

t.id AS tag\_id,

t.tag\_name,

COALESCE(SUM(pe.total\_engagement), 0) AS total\_engagement

FROM tags t

LEFT JOIN photo\_tags pt ON t.id = pt.tag\_id

LEFT JOIN PhotoEngagement pe ON pt.photo\_id = pe.photo\_id

GROUP BY t.id, t.tag\_name

)

SELECT

tag\_name,

total\_engagement

FROM HashtagEngagement

ORDER BY total\_engagement DESC;

* To determine the hashtags or content topics with the highest engagement rates, you can analyze data from social media platforms or use analytics tools to measure the performance of different posts.
* Once you have this information, you can use it to guide your content strategy and ad campaigns by focusing on creating more content around these high-performing hashtags or topics. This can help you tailor your messaging to what resonates with your audience the most, leading to higher engagement rates and better results for your campaigns.
* Additionally, you can use this data to identify trends and create timely and relevant content that is more likely to capture the attention of your audience. By understanding what topics or hashtags are driving the most engagement, you can optimize your content strategy to reach your target audience effectively and generate more meaningful interactions with your brand.

1. Are there any patterns or trends in user engagement based on demographics (age, location, gender) or posting times? How can these insights inform targeted marketing campaigns?

* The table highlights users with the highest engagement by month of account creation, since we lack demographic data. In 2016, accounts created in May, June, July, August, September, and December showed the most engagement, while in 2017, those created in January, February, and April stood out. This suggests that the optimal time for running targeted marketing campaigns would be either at the end of the year or at the beginning of the year.

SELECT id, username, account\_date, engagement

FROM (

SELECT u.id, u.username,

DATE\_FORMAT(u.created\_at, '%Y-%m') AS account\_date,

IFNULL(p.posts\_count, 0) + IFNULL(l.likes\_count, 0) + IFNULL(c.comments\_count, 0) AS engagement,

DENSE\_RANK() OVER(ORDER BY IFNULL(p.posts\_count, 0) + IFNULL(l.likes\_count, 0) + IFNULL(c.comments\_count, 0) DESC) AS ranking

FROM users u

LEFT JOIN (

SELECT user\_id, COUNT(id) AS posts\_count

FROM photos

GROUP BY 1

) p ON u.id = p.user\_id

LEFT JOIN (

SELECT user\_id, COUNT(user\_id) AS likes\_count

FROM likes

GROUP BY 1

) l ON u.id = l.user\_id

LEFT JOIN (

SELECT user\_id, COUNT(id) AS comments\_count

FROM comments

GROUP BY 1

) c ON u.id = c.user\_id

GROUP BY 1, 2, 3

) dt

WHERE ranking = 1;

1. Based on follower counts and engagement rates, which users would be ideal candidates for influencer marketing campaigns? How would you approach and collaborate with these influencers?

WITH followers AS(

SELECT u.id user\_id,

username,

COUNT(follower\_id) AS follower\_count

FROM users u

LEFT JOIN follows f ON u.id=f.followee\_id

GROUP BY 1,2

),

engagement AS(

SELECT u.id AS user\_id,

username,

p.id AS photo\_id,

(COUNT(DISTINCT c.user\_id)+ COUNT(DISTINCT l.user\_id)) AS engagement\_recieved

FROM users u

LEFT JOIN photos p ON u.id=p.user\_id

LEFT JOIN comments c ON c.photo\_id=p.id

LEFT JOIN likes l ON l.photo\_id=p.id

GROUP BY 1,2,3

),

influencers AS(

SELECT f.user\_id,

f.username,

follower\_count,

(SUM(engagement\_recieved)/follower\_count) avg\_engagement\_rate

FROM followers f

JOIN engagement e ON f.user\_id=e.user\_id

GROUP BY 1,2,3

)

SELECT \* FROM influencers

ORDER BY 3 DESC, 4 DESC

LIMIT 5;

* **Customized Outreach**:  
  Create personalized messages for each influencer by emphasizing their individual engagement levels and follower count. Recognize their content style and audience to show that we’ve done our homework. This helps to establish a sincere connection and sets a strong foundation for collaboration.
* **Campaigns Tailored to Niche**:  
  Design marketing campaigns that align with each influencer’s specific niche and audience. For example, if an influencer focuses on lifestyle content, we can develop campaigns that naturally integrate our products into their lifestyle posts. This makes the partnership feel more authentic and relatable to their followers, boosting the campaign’s impact.
* **Enticing Incentives**:  
  Provide compelling incentives for collaboration, such as competitive compensation, free products, or exclusive event access. We can also emphasize potential growth opportunities for influencers, like co-branded content or exposure to a wider audience through our platforms. This encourages influencers to actively promote our brand.
* **Clear Metrics and Regular Feedback**:  
  Define specific performance metrics to evaluate the success of influencer partnerships, such as tracking engagement, conversions, and reach. By conducting regular feedback sessions, we can refine future campaigns and strengthen relationships with influencers.
* **Ongoing Partnerships**:  
  Consider building long-term relationships with the top-performing influencers. This not only ensures consistent brand representation but also nurtures trust with their audience over time, making collaborations more meaningful and effective at driving brand awareness and loyalty.

1. Based on user behavior and engagement data, how would you segment the user base for targeted marketing campaigns or personalized recommendations?

WITH photoengagementCTE AS

(

SELECT

p.id AS photo\_id,

p.user\_id,

COUNT(DISTINCT l.user\_id) AS like\_count,

COUNT(DISTINCT c.id) AS comment\_count

FROM photos p

LEFT JOIN likes l ON p.id = l.photo\_id

LEFT JOIN comments c ON p.id = c.photo\_id

GROUP BY p.id

),

usersinvolveCTE AS

(

SELECT

pe.user\_id,

SUM(pe.like\_count) AS total\_likes,

SUM(pe.comment\_count) AS total\_comments,

SUM(pe.like\_count + pe.comment\_count) AS total\_engagement

FROM photoengagementCTE pe

GROUP BY pe.user\_id

)

SELECT

ue.user\_id,

u.username,

ue.total\_likes,

ue.total\_comments,

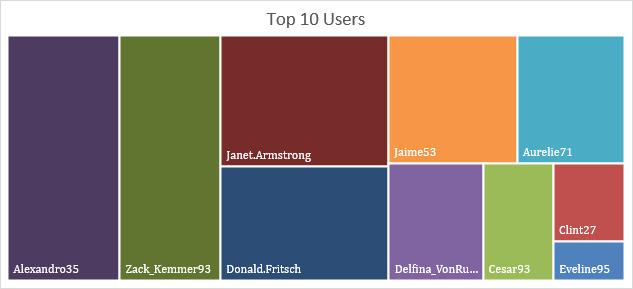
ue.total\_engagement,

rank() over (order by ue.total\_engagement desc) AS ranks

FROM usersinvolveCTE ue

JOIN users u ON ue.user\_id = u.id

ORDER BY ue.total\_engagement DESC;



To segment the user base for targeted marketing campaigns or personalized recommendations based on user behaviour and engagement data, you could consider the following segmentation criteria:

* Demographics: Segment users based on their age, gender, location, income level, education level, etc
* Psychographics: Segment users based on their interests, values, attitudes, and lifestyle choices.
* Behavioural data: Segment users based on their past interactions with your platform, such as pages visited, products purchased, frequency of visits, time spent on site, etc.
* Engagement level: Segment users based on their level of engagement with your platform, such as active users, dormant users, new users, etc.
* Purchase history: Segment users based on their past purchase history, such as high-value customers, frequent purchasers, one-time buyers, etc.

By analyzing user behaviour and engagement data using these segmentation criteria, you can create targeted marketing campaigns and personalized recommendations that resonate with specific user groups, ultimately driving higher engagement and conversions.

1. If data on ad campaigns (impressions, clicks, conversions) is available, how would you measure their effectiveness and optimize future campaigns?

If data on ad campaigns (impressions, clicks, conversions) is available, then the key metrics to measure effectiveness and optimize future campaigns would be –

* Click-Through Rate (CTR) = (Clicks / Impressions) \* 100
* Conversion Rate (CVR) = (Conversions / Clicks) \* 100
* Cost per Acquisition (CPA) = Total Spend / Conversions
* Return on Ad Spend (ROAS) = Revenue / Total Spend

For optimization of future campaign –

* Allocate budget to high performing segments.
* Adjust bids for better performing ads.
* Retargeting users who didn’t convert.
* Optimizing landing pages to improve conversion.

1. How can you use user activity data to identify potential brand ambassadors or advocates who could help promote Instagram's initiatives or events?

**Approach:**

1. **High Follower Count and High Engagement**:
   * Users with a high number of followers and consistently high numbers of likes and comments are prime candidates.
   * These users not only have a large reach but also have an engaged audience.
2. **Consistency in Engagement**:
   * Users who consistently receive high numbers of likes and comments across multiple posts may indicate genuine interest and influence, as opposed to one-time viral activity.
3. **Identify Users with Balanced Engagement**:
   * We should focus on users who both generate and receive engagement, as these users are more likely to influence their network and drive engagement for Instagram initiatives.

WITH followers AS(

SELECT u.id user\_id,

username,

COUNT(follower\_id) AS follower\_count

FROM users u

LEFT JOIN follows f ON u.id=f.followee\_id

GROUP BY 1,2

),

engagement AS(

SELECT u.id AS user\_id,

username,

p.id AS photo\_id,

COUNT(DISTINCT c.user\_id) comments,

COUNT(DISTINCT l.user\_id) likes

FROM users u

LEFT JOIN photos p ON u.id=p.user\_id

LEFT JOIN comments c ON c.photo\_id=p.id

LEFT JOIN likes l ON l.photo\_id=p.id

GROUP BY 1,2,3

)

SELECT f.user\_id,

f.username,

follower\_count,

SUM(comments) + SUM(likes) total\_likes

FROM followers f

JOIN engagement e ON f.user\_id=e.user\_id

WHERE follower\_count = (SELECT MAX(follower\_count) FROM followers)

GROUP BY 1,2,3

ORDER BY 4 DESC LIMIT 4;

1. How would you approach this problem, if the objective and subjective questions weren't given?

To address the objectives of increasing user engagement, retention, and acquisition, I would follow a systematic approach consisting of several steps:

**Assess User Engagement**  
**Objective**: Identify the most active users by evaluating their likes, comments, and photo uploads.  
• **Query 1**: Compute the total number of likes and comments for each user.

**Measure User Retention**  
**Objective**: Determine how long users have remained active on the platform and detect any retention trends.  
• Calculate the account age for each user based on the account creation timestamp.

**Analyze User Activity Trends**  
**Objective**: Investigate patterns in user activity over time to uncover engagement trends.  
• Track engagement metrics such as likes and comments over different time periods.

**Discover Potential Brand Advocates**  
**Objective**: Identify users who are highly engaged and have a large follower base, making them ideal candidates for brand ambassadorship.  
• Merge follower counts with engagement data to pinpoint potential brand advocates.

**Examine Tag Usage and Engagement**  
**Objective**: Analyze the most commonly used tags and their relationship to user engagement levels.  
• Determine the most popular tags and assess their corresponding engagement metrics.

Based on the insights gathered, I would like to suggest actionable recommendations to enhance user engagement, improve retention, and drive acquisition:

* **Utilize Potential Brand Ambassadors**:  
  Leverage users with high engagement and large follower counts for targeted marketing campaigns. Partner with these influencers to promote new features, events, or brand initiatives effectively.
* **Analyze Popular Content for Trends**:  
  Examine the photos with the most likes and comments to uncover user preferences. Apply these insights to shape content creation strategies, ensuring marketing campaigns resonate with user interests.
* **Enhance Discoverability Through Tags**:  
  Use findings from the tag usage analysis to boost discoverability. Identify trending tags and encourage users to include them in their posts to increase visibility.
* **Gather Feedback to Address Disengagement**:  
  Set up feedback systems to understand why some users become inactive. Combine retention metrics with user feedback to gain insights into areas where improvements can be made.
* **Incorporate Gamification to Drive Interaction**:  
  Implement gamification techniques to boost user engagement. Offer rewards for actions like posting, commenting, liking, and using specific tags.
* **Regularly Analyze Data for Emerging Trends**:  
  Continuously review engagement, retention, and acquisition metrics to spot new trends and adjust marketing strategies as needed to stay relevant.

1. Assuming there's a "User\_Interactions" table tracking user engagements, how can you update the "Engagement\_Type" column to change all instances of "Like" to "Heart" to align with Instagram's terminology?

UPDATE User\_Interactions

SET Engagement\_Type = 'Heart'

WHERE Engagement\_Type = 'Like';

* Update the Engagement\_Type column in the User\_Interactions table to change all instances of "Like" to "Heart", we can use an SQL UPDATE statement.