**Learners have to come up with a Report to support the answers to the following questions and suggestions**

Objective Questions

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

Ans: There are no duplicate values in any tables.

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

Ans:

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).

Ans:

select p.id as PostId,avg(pt.tag\_id) as AvgPerPost

from photos p

join photo\_tags pt on p.id=pt.photo\_id

group by p.id

order by p.id

* Using join Photos and Photo\_tags table on photo\_id,calculate average no.of tags per post using avg() aggregate function and grouped them using group by function.

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

Ans:

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;

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?

Ans:

WITH followers\_count AS (

SELECT follower\_id,COUNT(follower\_id) AS num\_followers

FROM follows

group by follower\_id

),

followings\_count AS (

SELECT followee\_id,COUNT(followee\_id) AS num\_followings

FROM follows

group by followee\_id

)

SELECT u.username,

MAX(f.num\_followers) AS max\_followers,

MAX(f1.num\_followings) AS max\_followings

FROM users u

left JOIN followers\_count f ON u.id = f.follower\_id

left JOIN followings\_count f1 ON u.id = f1.followee\_id

GROUP BY u.username

ORDER BY max\_followers DESC, max\_followings DESC

Created 2 different CTEs to count followers and followee using using follow and user table.

I used user table and both cte in select query using left join to show maximum followers and followee for each user.

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

Ans:

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,

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

),

involveduserCTE as

(

select

pe.user\_id,

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

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

from photoengagementCTE 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 involveduserCTE 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)

Ans:

select id,username

from users where id not in (select user\_id from likes)

I used users table to display details of users.

* Used not in to get details of user who never liked any post.

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

Ans:

To leverage user-generated content for personalized and engaging ad campaigns, you can use the data from the tables mentioned above in the following ways:

1. Analyze the comments and likes on photos to understand user preferences and interests. This data can help you identify popular trends and topics among your users.

2. Use the follow relationships between users to target ads to users who are likely to be interested in similar content. For example, if a user follows another user who frequently posts about fashion, you can target ads related to fashion to both users.

3. Utilize the tags associated with photos to categorize content and target ads based on specific interests. For example, if a user frequently interacts with photos tagged with #travel, you can show ads related to travel destinations or products.

4. Collaborate with influencers or users with a large following to create user-generated content that can be used in ad campaigns. This type of content is often more authentic and engaging for users.

Overall, by leveraging user-generated content and user data from your database, you can create personalized and engaging ad campaigns that resonate with your target audience.

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?

Ans:

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.

This information can guide content creation and curation strategies by helping creators focus on producing more of the types of content that generate higher levels of engagement. For example, if users are consistently engaging with photos more than videos, creators may want to prioritize creating more visual content. Alternatively, if reels are driving the most engagement, creators may want to dedicate more resources to producing short-form video content.

By understanding which content types are most popular with their audience, content creators can optimize their content strategy to maximize engagement, reach, and overall success on their platform. Additionally, by regularly monitoring user activity data and adjusting content strategies accordingly, creators can stay relevant and continue to attract and retain a loyal following.

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

Ans:

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.

Ans:

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.

In out data,dates in comments,photos and likes table are same and of last month.

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.

Ans:

with avg\_likes\_per\_tag as

(

select

pt.tag\_id,

round(avg(l.user\_id),2) as avg\_likes

from photo\_tags pt

join photos p on pt.photo\_id = p.id

join likes l on p.id = l.photo\_id

group by pt.tag\_id

)

select

t.tag\_name,

alt.avg\_likes

from avg\_likes\_per\_tag alt

join tags t on alt.tag\_id = t.id

order by alt.avg\_likes desc;

As it is already mentioned in question,created CTE to calculate average likes of posts.

And in select query displayed that tag with their average like in descending to order to see highest to lowest.

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

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

select

distinct u1.username

from users u1

join follows f1 on u1.id = f1.follower\_id

join follows f2 on u1.id = f2.followee\_id

where f1.created\_at < f2.created\_at;

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?

Ans:

The most loyal and valuable users are those who consistently engage with the platform, contribute high-quality content, participate in discussions, and positively impact the community. These users are likely to be active members who regularly provide feedback, support other users, and demonstrate a strong commitment to the platform.

To reward and incentivize these loyal users, you could consider implementing a loyalty program that offers various benefits and perks such as exclusive access to premium features, discounts on products or services, early access to new features, recognition in the community, badges or titles to display on their profile, or even physical rewards like merchandise or gift cards. Additionally, you could create special events or challenges for these users to participate in, creating a sense of exclusivity and importance within the community.

Overall, the key is to show appreciation for their continued support and engagement, and to provide incentives that are meaningful and relevant to their interests and motivations. By recognizing and rewarding these loyal users, you can encourage them to remain active, engaged, and committed to the platform in the long run.

with LikesGiven as

(

select user\_id, COUNT(\*) as likes\_given

from likes

where created\_at >= NOW() - interval 1 month

group by user\_id

),

LikesReceived 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

),

CommentsGiven as

(

select user\_id, COUNT(\*) as comments\_given

from comments

where created\_at >= NOW() - interval 1 month

group by user\_id

),

CommentsReceived 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

),

PhotosUploaded as

(

select user\_id, COUNT(\*) as photos\_uploaded

from photos

where created\_dat >= NOW() - interval 1 month

group by user\_id

),

Followers as

(

select followee\_id as user\_id, COUNT(\*) as followers

from follows

group by followee\_id

),

Followings 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 LikesGiven lg on u.id = lg.user\_id

left join LikesReceived lr on u.id = lr.user\_id

left join CommentsGiven cg on u.id = cg.user\_id

left join CommentsReceived cr on u.id = cr.user\_id

left join PhotosUploaded pu on u.id = pu.user\_id

left join Followers f on u.id = f.user\_id

left join Followings 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;

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

Ans:

According to our database, all users are active users. We didn’t find any inactive user for last 3 months using below query

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;

For recommendations:

1. Personalized outreach: Send personalized messages to inactive users thanking them for being part of the community in the past and encourage them to start posting again. Ask for feedback on why they are no longer active and what could be done to improve their experience.

2. Exclusive offers or incentives: Offer special promotions or incentives for inactive users to encourage them to start posting or engaging again. This could include discounts, freebies, or exclusive access to premium features.

3. Create a sense of urgency: Use limited-time offers or time-sensitive promotions to create a sense of urgency and encourage inactive users to start engaging again.

4. Highlight new content or features: Showcase any new content or features that have been added to the platform since the user was last active. This can pique their interest and encourage them to start posting again.

5. Organize contests or challenges: Organize contests or challenges that require active participation from users. This can help re-engage inactive users and encourage them to start posting or engaging again.

6. Provide relevant and engaging content: Make sure that the content on the platform is relevant, engaging, and interesting to the target audience. This can help re-engage inactive users and encourage them to start posting again.

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

Ans:

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

For Strategies:

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?

Ans:

Insights:

There can be patterns and trends in user engagement based on demographics and posting times. For example, younger audiences may be more active on social media platforms during evenings, while older audiences may be more active during the daytime. Also, users in different locations may have different preferences in terms of content and engagement.

These insights can inform targeted marketing campaigns by helping marketers understand when and how to reach their target audience. For example, if a brand is targeting a younger demographic, they may want to schedule posts in the evenings when they are most likely to be active. Similarly, if a brand is targeting a specific location, they can tailor their content to appeal to the preferences and interests of that particular audience.

By analyzing these patterns and trends, marketers can create targeted campaigns that are more likely to resonate with their intended audience, leading to higher engagement and ultimately, conversions. This data can also help marketers optimize their advertising spend by focusing on the times and locations where their target audience is most active.

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?

Ans:

Ideal candidates for influencer marketing campaigns would be users who have a high number of followers and a high engagement rate on their posts. This indicates that their audience is actively engaged and interested in their content, making them more likely to be receptive to promotional messages.

To approach and collaborate with these influencers, you could start by reaching out to them through direct messages on social media platforms or through their business email if available. In your message, be sure to introduce yourself and your brand, explain why you think they would be a good fit for your campaign, and outline the collaboration details such as compensation, deliverables, and timeline.

When collaborating with influencers, it's important to establish clear expectations and guidelines for the campaign to ensure it aligns with your brand values and objectives. Provide them with creative freedom to showcase your product or service in a way that resonates with their audience, while also ensuring that they disclose any sponsored content in accordance with FTC guidelines.

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

Ans:

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;

Insights:

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:

1. Demographics: Segment users based on their age, gender, location, income level, education level, etc

2. Psychographics: Segment users based on their interests, values, attitudes, and lifestyle choices.

3. 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.

4. Engagement level: Segment users based on their level of engagement with your platform, such as active users, dormant users, new users, etc.

5. 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?

Ans:

There are several ways to measure the effectiveness of ad campaigns and optimize future campaigns based on the data available:

1. Calculate the click-through rate (CTR) of the ad campaign by dividing the number of clicks by the number of impressions. A high CTR indicates that the ad is engaging and resonating with the audience.

2. Measure the conversion rate by dividing the number of conversions (e.g., purchases, sign-ups) by the number of clicks. This will help determine how effective the ad is at driving actions from users.

3. Conduct A/B testing by creating variations of the ad (e.g., different images, headlines, call-to-actions) and testing them against each other to see which performs best. This will help optimize future campaigns by identifying what resonates with the audience.

4. Utilize audience targeting options to reach specific demographics, interests, and behaviors that are most likely to convert. This will help improve campaign effectiveness by reaching the right audience.

By continuously measuring, analyzing, and optimizing ad campaigns based on these metrics, marketers can improve the effectiveness of their campaigns and achieve better results in the future.

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

Ans:

One way to use user activity data to identify potential brand ambassadors or advocates on Instagram is to analyse their engagement metrics. Look for users who consistently like, comment, and share posts related to Instagram's initiatives or events. These individuals are likely to be passionate about the brand and are more likely to promote it to their own followers.

Additionally, you can track the number of followers and reach of these users to identify those with a larger influence and greater potential to reach a wider audience. You can also analyse the type of content they typically post and their level of engagement with their own followers to ensure they align with the brand's values and messaging.

By identifying and reaching out to these potential brand ambassadors or advocates, you can collaborate with them to promote Instagram's initiatives or events, leveraging their influence and credibility to reach a larger audience and drive engagement.

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

Ans:

**Project Description: -**

Imagine you're a data analyst working with the product team at Instagram. Your role involves analysing 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, the product team might use them to decide on new features to build, and the development team might use them to improve the overall user experience. In this project, you'll be using SQL and MySQL Workbench as your tool to analyse Instagram user data and answer questions posed by the management team. Your insights will help the product manager and the rest of the team make informed decisions about the future direction of the Instagram app. Remember, the goal of this project is to use your SQL skills to extract meaningful insights from the data. Your findings could potentially influence the future development of one of the world's most popular social media platforms.

**SQL Tasks: -**

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

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.

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.

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 commonly used hashtags on the platform.

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:

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.

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: -**

For this project, I have used My SQL to extract the required data from the given database using the Join function, subqueries, Aggregation, where condition, Group by, Distinct and other functions required. keeping the Primary key and foreign key in consideration provided all the reports asked by the marketing department and Investor metrics department.

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?

Ans:

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.

UPDATE User\_Interactions

SET Engagement\_Type = 'Heart'

WHERE Engagement\_Type = 'Like';