**Social Media Analysis (SQL Project)  
  
 - Shreyas K**



**Data Description**

● comments\_id : unique identifier for each comment

● comment\_text : text content of a given comment

● user\_id : unique identifier for each user

● photo\_id : unique identifier for each photo

● created\_at : date of interaction in the form like, photos, tags

● follower\_id : user\_id of the follower for a certain user

● followee\_id : user\_id of followee for a certain user

● tag\_id : unique identifier for each tag

● image\_url : link to the image posted on the platform

● username : username chosen by the user

**Problem statement**

You are hired as a data analyst at Meta and asked to collaborate with Marketing team. Marketing teams wants 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**

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

**Ans-**

**For checking duplicate or missing NULL values below queries can be used. And this same process can be used for different tables of dataset.**

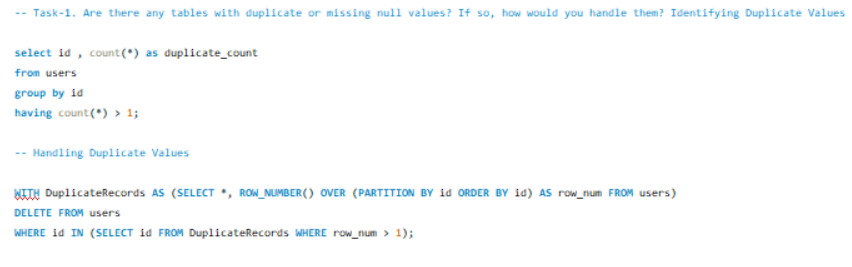
**After going through each table it is found that there is no duplicate or missing null values in the table.**

**Identifying Duplicate Values-**

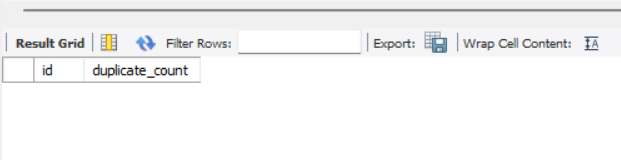
To find duplicates, I used a query that groups by the unique column “id” and uses a **HAVING** clause to identify records with more than one occurrence

**Handling Duplicate Values –**

To delete duplicates values, I use a **DELETE** query with a **ROW\_NUMBER()** window function , it will delete the values which have instances more than **1**.

****

**Result for duplicate values-**

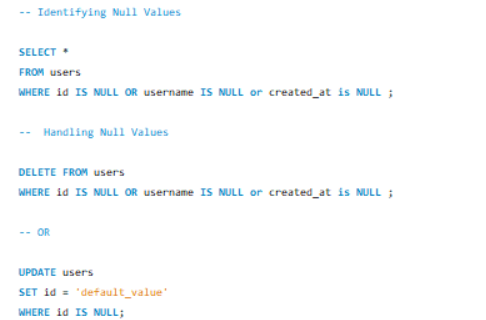
****

**Identifying Null Values-**

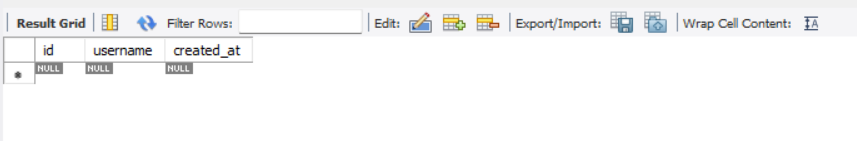
To find rows with null values I use **WHERE** clause that checks for nulls.

**Handling** **Null Values-**

If **NULL** values are not required for analysis then they can be deleted or if they are required for the analysis then they are updated with **default value** .



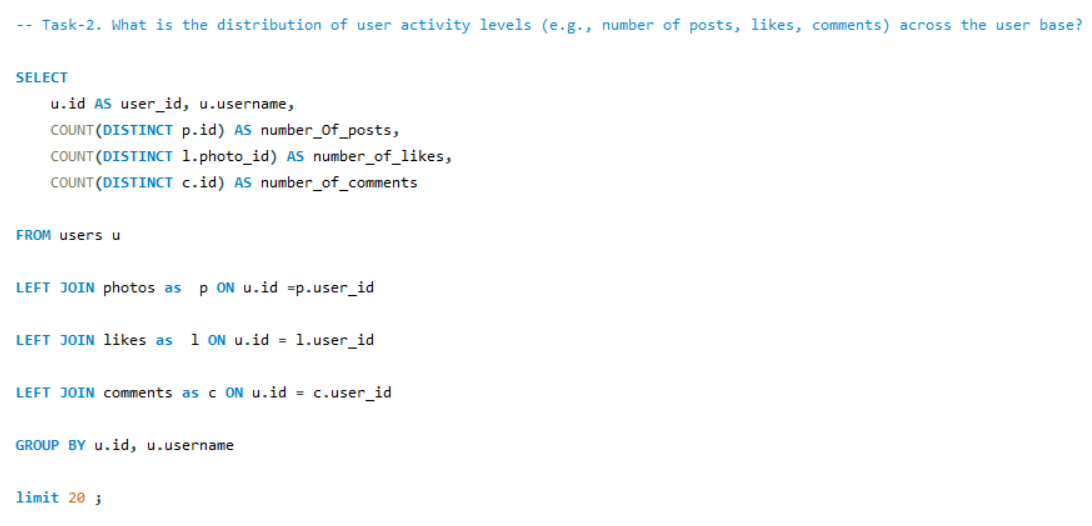
**Result for NULL values -**

****

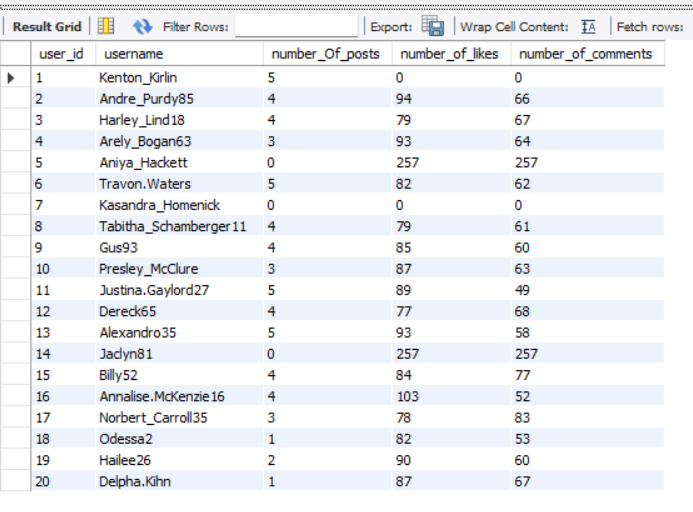
**Task-2. What is the distribution of user activity levels (e.g., number of posts, likes, comments) across the user base?**

**Ans –**

The below query retrieves a summary of user activity for the first 20 users, including the number of posts, likes, and comments each user has made.The query uses **LEFT JOIN** to ensure that all users from the users table are included in the result, even if they have no posts, likes, or comments. It count unique posts, likes, and comments for each user. The query groups results by each user, providing insights across users. The **`LIMIT 20`** restricts the output to 20 users.

****

**Result -**

****

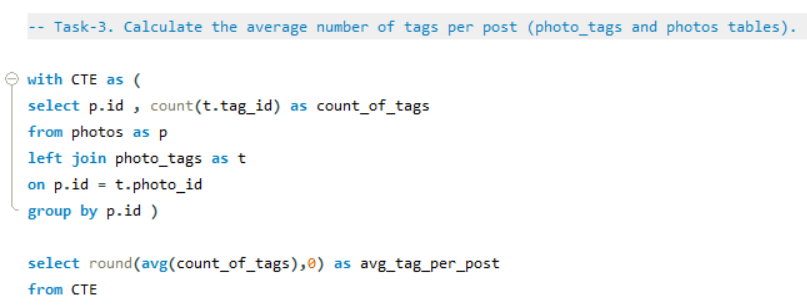
**Graph of Activity level-**

**Task-3. Calculate the average number of tags per post (photo\_tags and photos tables).**

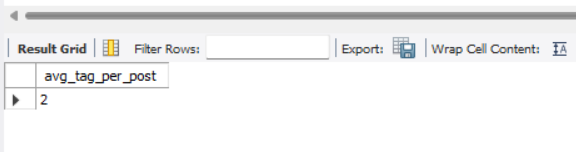
**Ans –**

In this query **Common Tabel Expression (CTE)** is used which form a temporary table for further use and will improve the readability of query. Here I calculated the count of tags by grouping on posts. **LEFT JOIN** is used to join photo\_tags and photos tables (Left join is use as it will keep the post which have 0 tags also , by this we will get the correct average of tags).

The **average**  is calculated on count of tags and round off to near number as tags can’t be in decimals. This is how I calculated the average tags for each post.

****

**Result -**

****

**Task-4. Identify the top users with the highest engagement rates (likes, comments) on their posts and rank them.**

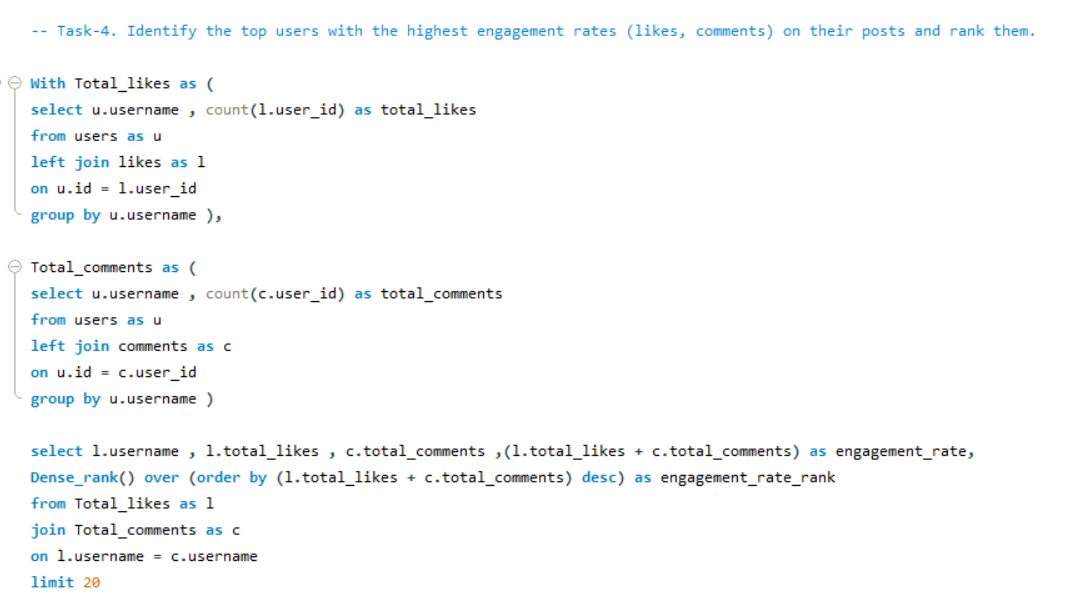
**Ans –**

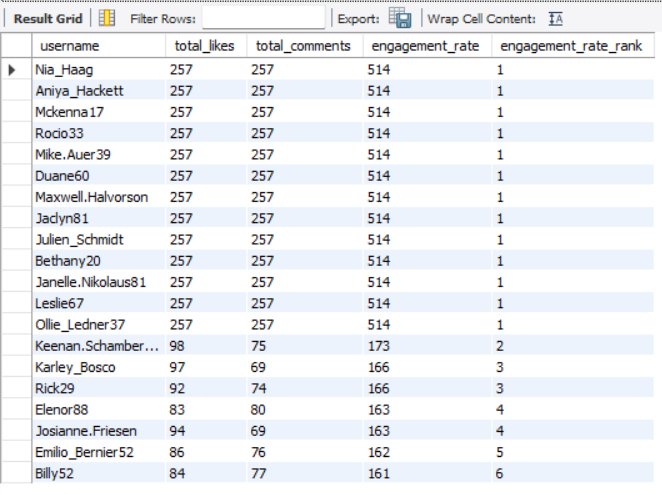
1.First I create a **Common Table Expression (CTE)** called Total\_likes to calculate the total likes each user has received. This is done by joining the users and likes tables and counting likes per user.

2.For the Total\_likes CTE, the **LEFT JOIN** between users and likes means that every user will appear in the result, even if they have no likes. If a user has no likes, total\_likes will simply show as 0 for that user.

3.Similarly , 2nd CTE Total\_comments is formed to calculate total comments by joining users and comments table on **u.id = c.user\_id** . This will give total\_likes and total\_comments. We can add them to get Engagement rate.

4.A **dense rank** is applied to rank users based on their engagement rate, with the highest engagement rate receiving rank 1. Finally, it limits the results to the **top 20 users** to show the results properly.



**Result –**

**Graph for Top Users-**

**Task-5. Which users have the highest number of followers and followings?**

**Ans –**

1.**Common Table Expression (CTE)** Count\_of\_followers ,counts the number of followers for each user by **grouping** follower\_id in the **follows** table.

2.Another CTE, **Count\_of\_followee**, counts the number of followings for each user.

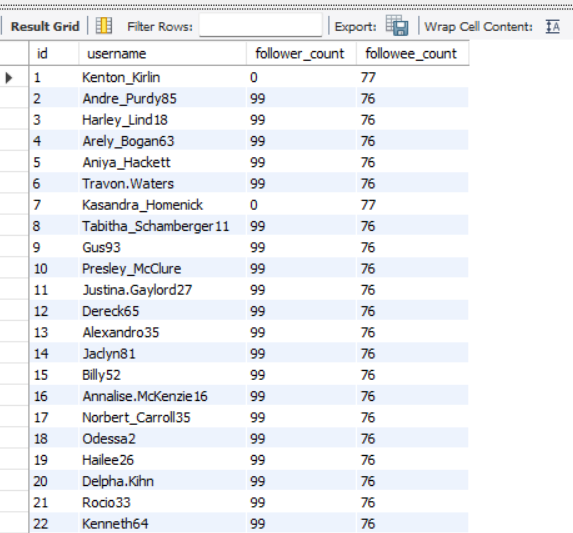
3. In the main query, it selects each user’s id and username from the users table.

4. **LEFT JOIN** to join the users with the follower and followee counts from CTEs, ensuring each user appears even if they have no followers or followings.

5. **COALESCE** is applied to replace **NULL** values with **0** for users with no followers or followings.



**Result -**

****

**Task-6.** **Calculate the average engagement rate (likes, comments) per post for each user.**

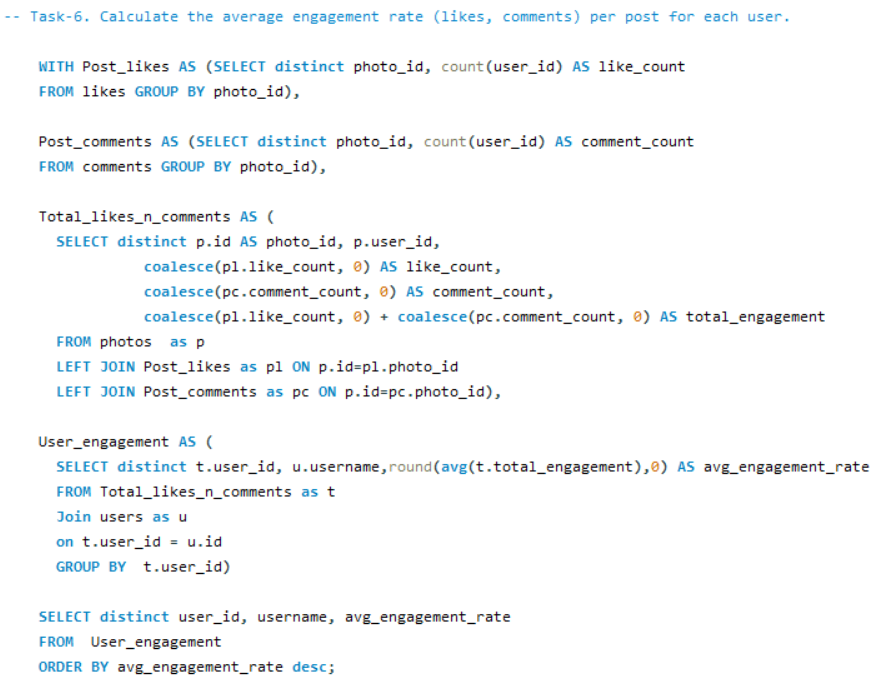
**Ans –**

**1.Post\_likes CTE:** We first calculate the total number of likes by grouping photo\_id in likes tabel.

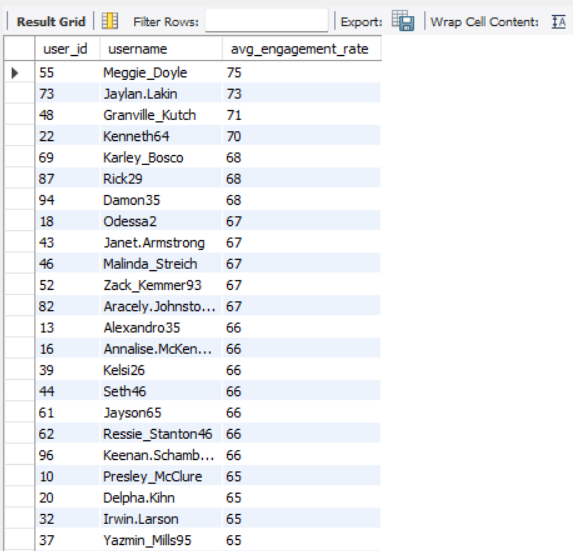
**2.Post\_comments CTE:** Similarly, we count the total comments per photo by grouping photo\_id in the comments table.

**3.Total\_likes\_n\_comments CTE:** Here, we join the photos table with Post\_likes and Post\_comments to get each photo’s like and comment count, setting missing values to zero with coalesce. We also calculate the sum of likes and comments for each photo as total\_engagement .

**4.User\_engagement CTE**: Next, we group by each user to calculate their average engagement rate per post, using avg(total\_engagement) and round it to near number.

**5.Final Query: Lastly,** we retrieve user\_id, username, and avg\_engagement\_rate from User\_engagement and sort by avg\_engagement\_rate in descending order.

**Result -**



**Graph for Avg\_Engagement\_rate-**

**Task-7. Get the list of users who have never liked any post (users and likes tables).**

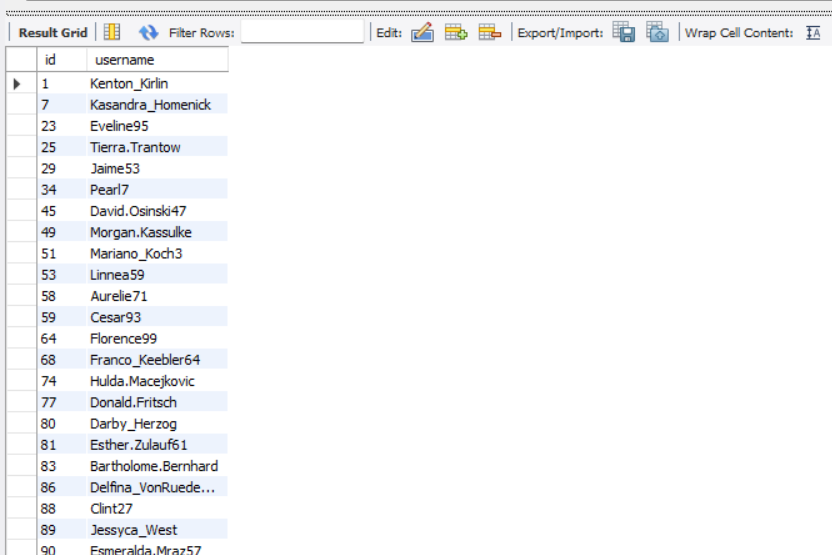
**Ans-**

1.In the **main select** statement, we took id and username from the users table.

2.We use a **where** clause to filter out any id values that appear in a subquery, which selects all user\_ids from the likes table.

****3.In this way, we will get users whose id doesn’t exist in the likes table(those who have never liked a post).

**Result -**



**Task-8. How can you leverage user-generated content (posts, hashtags, photo tags) to create more personalised and engaging ad campaigns?**

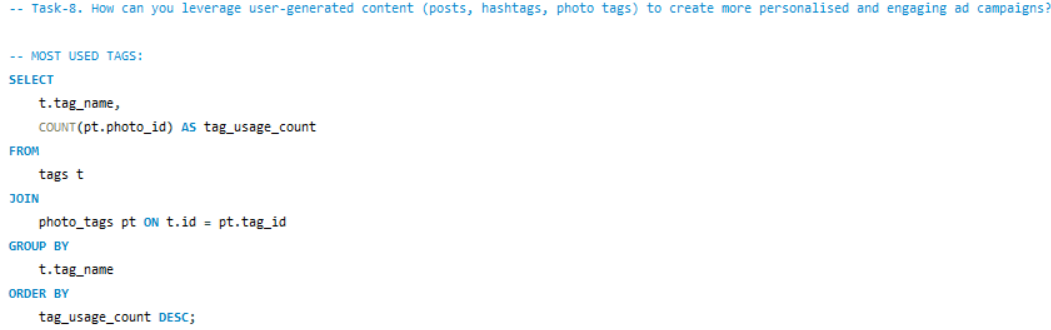
**Ans –**

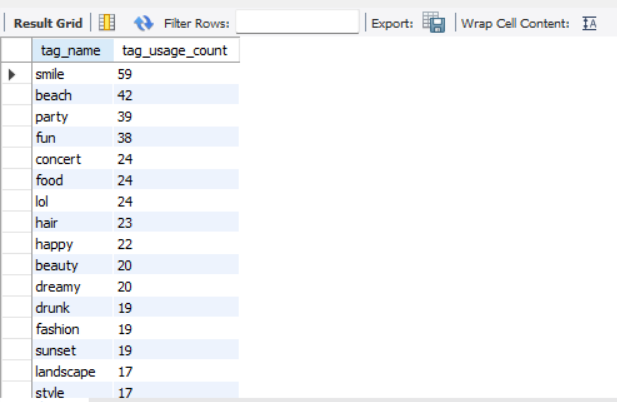
1.This query calculates the average number of likes for photos associated with each tag.

2.It first computes the total likes per photo and then averages these likes across all photos tagged with each tag.

3.The results are grouped by tag name and sorted in descending order of average likes.

4.Finally, the top 10 tags with the highest average likes are selected and displayed.

****

**Result –**

1.This query counts how often each tag is used in photos.

2.It joins the tags and photo\_tags tables to link tags with photos and then groups the results by tag name.

3.The query orders the tags by their usage count in descending order, showing which tags are most frequently associated with photos.

4.The final output displays each tag's name and the number of times it has been used.

****

**Suggestions-**

**1.Analyze User Interests**: I would start by identifying popular hashtags and photo tags to see what themes or topics users are engaging with. This insight allows us to create ads that resonate with these themes, increase engagement.

**2.Target Specific Audiences:** I’d use data from posts and tags to group users into segments based on their interests and behaviors. This lets us target ads specifically to these groups, increasing the chance they’ll engage with the content.

**3.Leverage Social Proof:** Featuring user comments, or popular posts in campaigns can build trust and drive conversions. This showcases real user experiences and encourages others to participate, boosting ad effectiveness.

**4.Promote Influencers and Power Users:** Engaging with users who are particularly active with relevant tags orcontent can amplify brand reach. Brands can offer these users special incentives or collaborations, letting their authentic voice enhance brand credibility and expand its reach.

**5.Keep Ads Updated with Trends:** By tracking trending hashtags, I could help keep ads aligned with current trends, ensuring they’re fresh and engaging for the audience.

**Task-9. 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 –**

**Method for Determining Correlation:**

1. To determine whether user activity levels and content kinds are related, I would begin with a correlation analysis utilizing statistical techniques. This would let us to determine whether users are more interested in reels, videos, or photographs.

2. If there’s a strong correlation between high engagement and certain content types, it could guide us in focusing more on those types to maximize user interaction and satisfaction.

**Guiding Content Creation and Curation:**  
  
**1.Include Trendy subjects:** Using popular hashtags and trendy subjects on a frequent basis will assist provide new and pertinent content to keep visitors interested.  
  
**2.Customized Suggestions:** We may make recommendations for material that corresponds with what users most frequently watch or follow based on their interests and previous interactions.  
  
**3.Use of Relevant Hashtags:** Adding hashtags associated with trending subjects might help increase user engagement because they are likely to draw in visitors and promote conversation.

**Task-10. Calculate the total number of likes, comments, and photo tags for each user.**

**Ans –**

1.The likes\_count **Common Table Expression (CTE)** counts the number of likes each user has received, by groups them by user\_id.

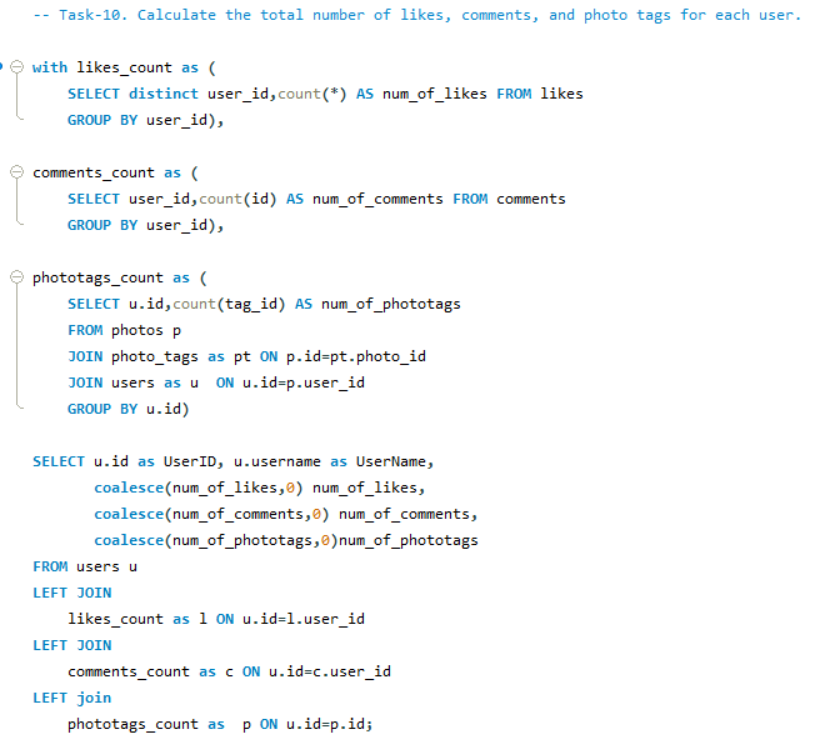
2. The comments\_count CTE counts the total number of comments each user has made.

3.The phototags\_count CTE counts the number of times each user’s photos have been tagged. This is done by joining the photos, photo\_tags, and users tables, linking photos to tags and photos to their respective users, then grouping by user ID.

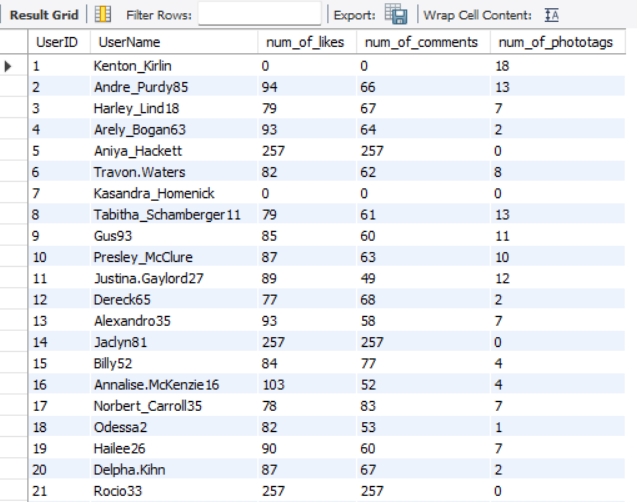
4. The **main query** selects each user’s ID and username from the users table. It left joins with each CTE to bring in the counts of likes, comments, and photo tags, matching on user\_id.

5. The query uses **COALESCE** to replace **NULL** values with zero, for each metric (likes, comments, and photo tags).

6. The query returns each user’s ID (UserID), username (UserName), and their respective counts for likes (num\_of\_likes), comments (num\_of\_comments), and photo tags (num\_of\_phototags).

****

**Result-**

****

**Graph for Number of likes,comments and tags-**

**Task-11. Rank users based on their total engagement (likes, comments, shares) over a month.**

**Ans –**

**The data of likes and comments is given only for one month i.e. 11th month and year is 2024.**

1.The Post\_likes **Common Table Expression (CTE)** count the total number of likes each user received, and stores this count as like\_count.

2.The Post\_comments CTE groups data by user\_id and count the total number of comments each user made, and stores this count as comment\_count.

3.The Total\_likes\_n\_comments CTE combines the data from Post\_likes and Post\_comments by joining them with the users table on user\_id. This fetch each user’s username, like\_count, and comment\_count.

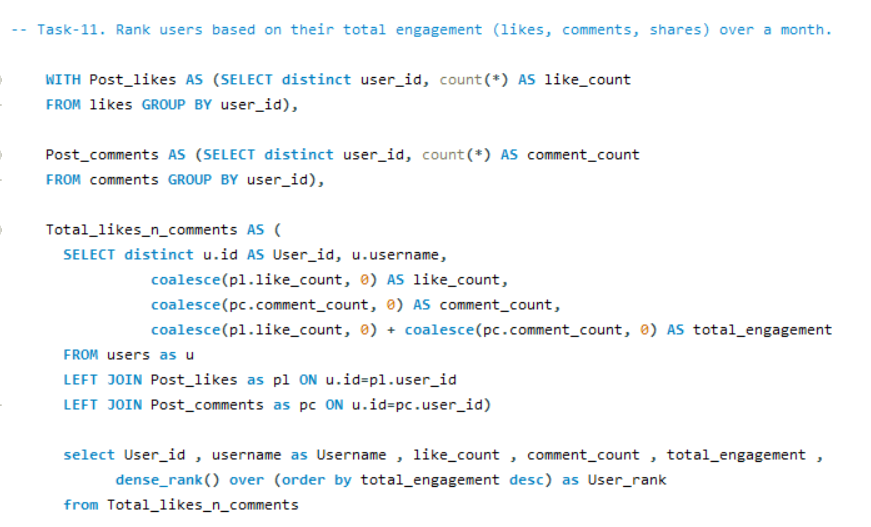
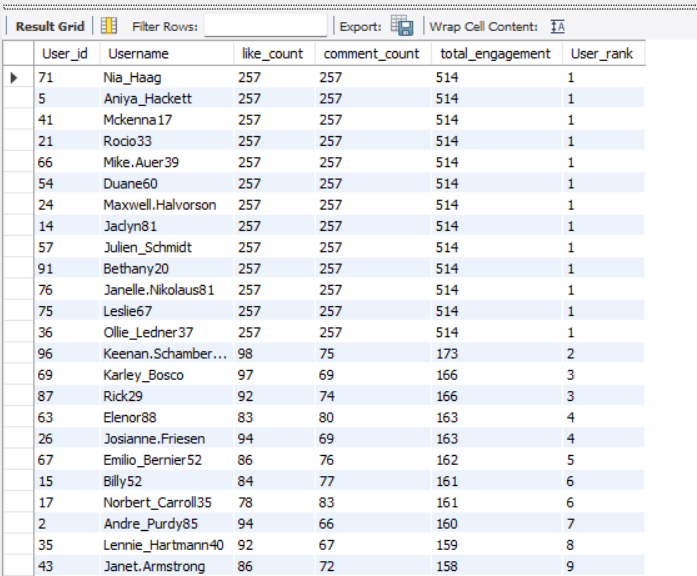
4. **COALESCE** function is used to handle null values, replacing any missing like or comment counts with zeroes.

5.A total\_engagement column is created by adding like\_count and comment\_count for each user. This column represents the overall engagement for each user.

6.The main **SELECT** statement retrieves User\_id, Username, like\_count, comment\_count, and total\_engagement from the Total\_likes\_n\_comments CTE.

7.It also uses the **DENSE\_RANK()** window function to rank users based on total\_engagement in descending order, with ties getting the same rank.

8.In this way query provides a ranked list of users by engagement, showing each user’s ID, username, number of likes and comments, total engagement, and engagement rank.

**Result-**

**Graph for Total Engagement –**

**Task-12. 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-**

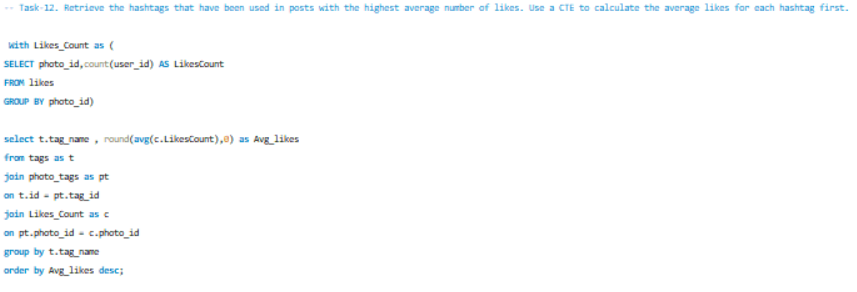
1.The **Likes\_Count CTE** count the number of likes each photo has received and label it as LikesCount.

2. In the main query, the tags table (which contains hashtag information) is **joined** with the photo\_tags table,it link hashtags to photos via tag\_id and photo\_id.

3. The query then joins the Likes\_Count CTE to get the like counts for each photo linked to specific hashtags.

4. For each hashtag (tag\_name), the **AVG function** calculates the average of LikesCount and rounds this average to zero decimal places, labeling it as Avg\_likes.

5. The query groups the results by tag\_name to ensure each hashtag appears once with its average likes, then orders the results by Avg\_likes in **descending** order to show hashtags with the highest average likes at the top.

****

**Result -**

****

**Graph for Avg\_likes\_for\_hashtag –**

**Task-13. Retrieve the users who have started following someone after being followed by that person.**

**Ans –**

**Followed\_User:** The user who was initially followed.

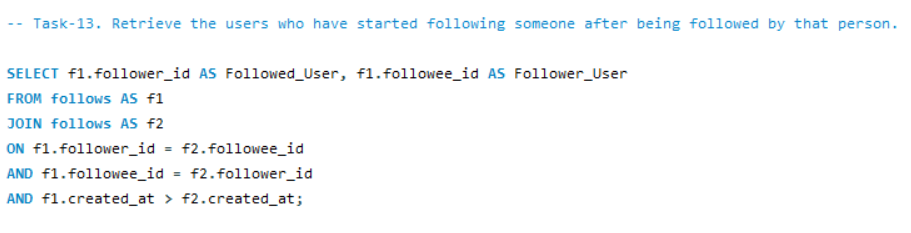
**Follower\_User:** The user who followed back after being followed.

1.I perform a self-join on the follows table, treating one instance of the table as f1 (representing one follow action) and the other as f2 (representing the another follow action).

2.Later match f1.follower\_id with f2.followee\_id to find the user who was followed and f1.followee\_id with f2.follower\_id to find the user who initiated the follow action back.

3.The condition **f1.created\_date > f2.created\_date** is added to ensure that the follow-back (f1) happened after the initial follow (f2).

**No record found for the users who have started following someone after being followed by that person.**

**Result-**

****

**subjective questions**

**Task-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 query calculates user engagement metrics on a platform, including likes, comments, photos, photo tags, and follower counts.**

**Common Table Expressions (CTEs):**

1.likes\_count: Counts likes per user.

2.comments\_count: Counts comments per user.

3.photo\_counts: Counts photos per user.

4.phototags\_count: Counts photo tags per user.

5.Count\_of\_followers: Counts followers per user.

6. Joins each **CTE** with the users table using **LEFT JOIN** to include all users, even those without specific activities.

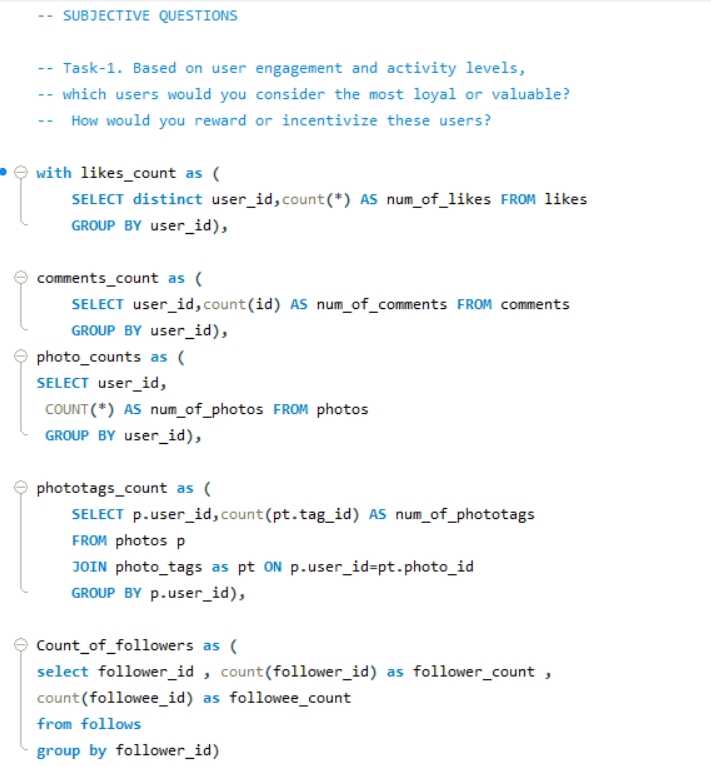
7.Uses **COALESCE** to replace **NULL** values with zero.

8. Calculated an engagement\_rate as the sum of likes, comments, photos .

9.Then ranked users by engagement\_rate in descending order using **DENSE\_RANK().**

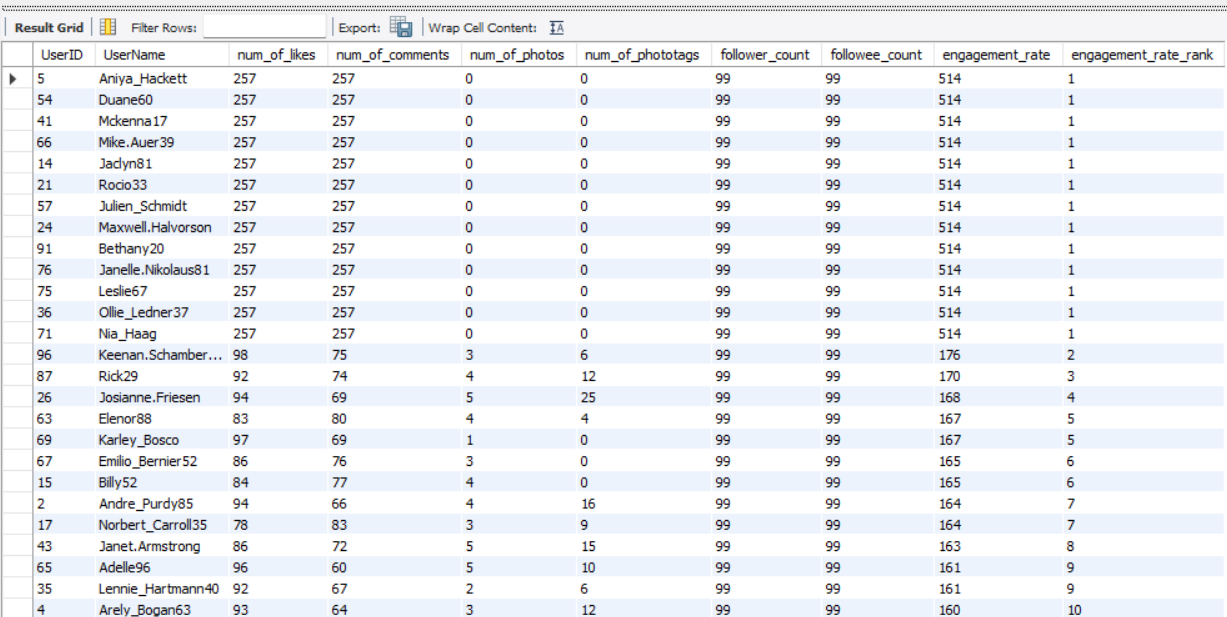
10.This ranking will give us the users with high engagement\_rate.

**This is how we will rank users based on user engagement and activity levels. The highly rank users can be considered as the loyal or valuable.**

****

****

**Result –**

****

**Ways for rewarding and incentivizing the users:**

**1.Invite them to events:**  By inviting them to participate in events organized by the platform, you can develop a stronger sense of community among them and help them to expand their network.

**2.By giving discounts and gift card :** This is one of the most popular yet effective ways to honour a valued user. Because they find these rewards exciting, this will encourage them to utilize the site more.

**3. Custom titles or badges:** This helps them to stand out from the other users and this will make them feel valued.

**4. Easy access to specific tools or features:** This is one method of assisting the valued clients in encouraging their expansion and streamlining the procedure. Giving them access to analytics and engagement insights will improve their comprehension of their users.

**5.Giving them the first access :** The access of newly launch feature can be first given to the loyal or valued users. This will make them feel valued and with can promote this feature as they have great engagement rate. This will help company also.

**Task-2. For inactive users, what strategies would you recommend to re-engage them and encourage them to start posting or engaging again?**

**Ans –**

**The query calculates user engagement metrics on a platform, including likes, comments, photos, photo tags, and follower counts.**

**Common Table Expressions (CTEs):**

1.likes\_count: Counts likes per user.

2.comments\_count: Counts comments per user.

3.photo\_counts: Counts photos per user.

4.phototags\_count: Counts photo tags per user.

5.Count\_of\_followers: Counts followers per user.

6. Joins each **CTE** with the users table using **LEFT JOIN** to include all users, even those without specific activities.

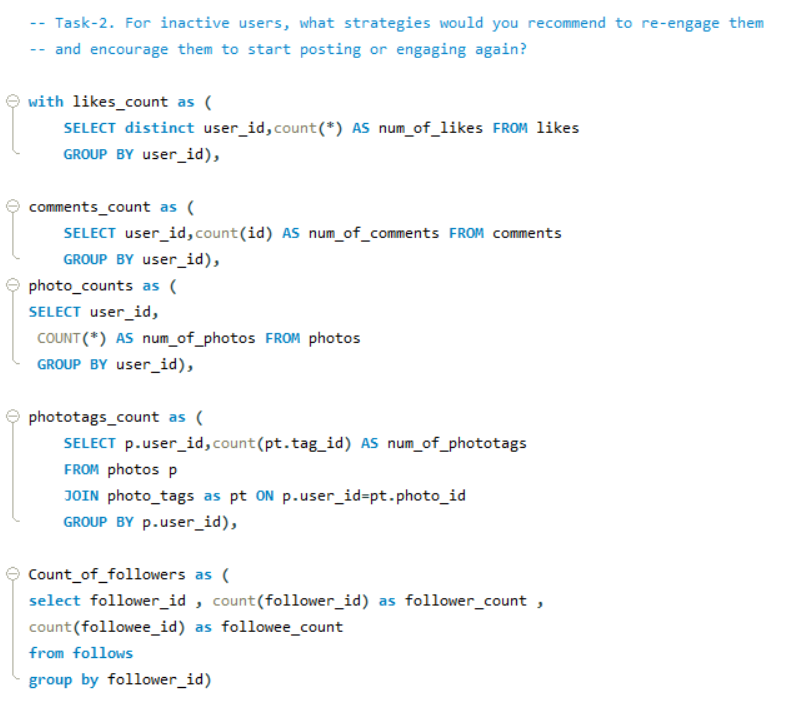
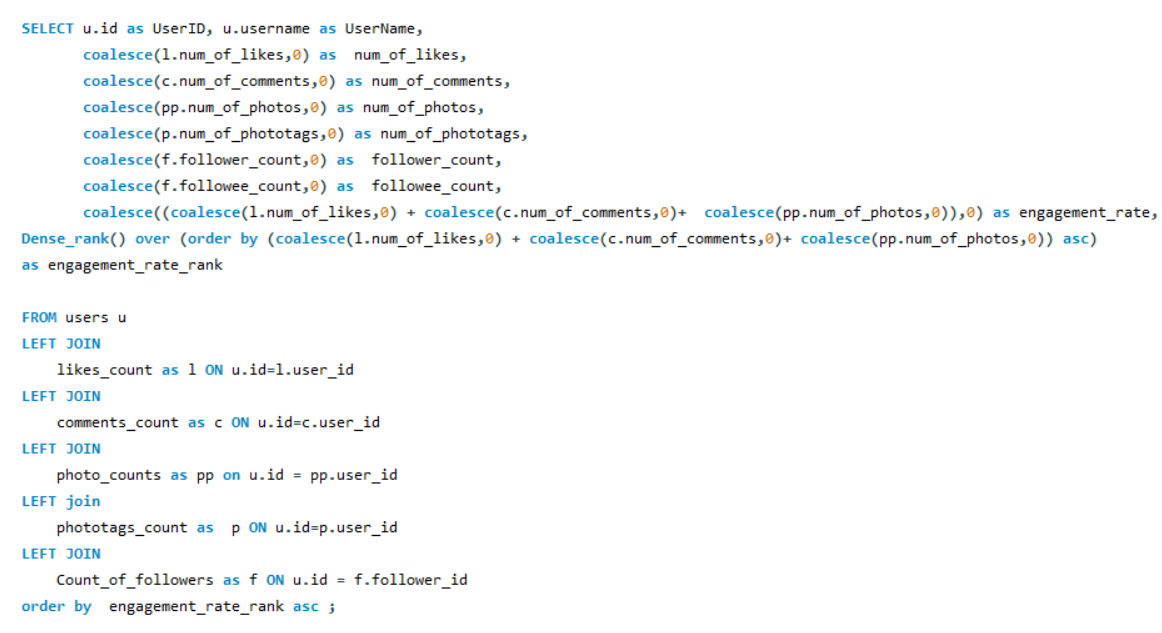
7.Uses **COALESCE** to replace **NULL** values with zero.

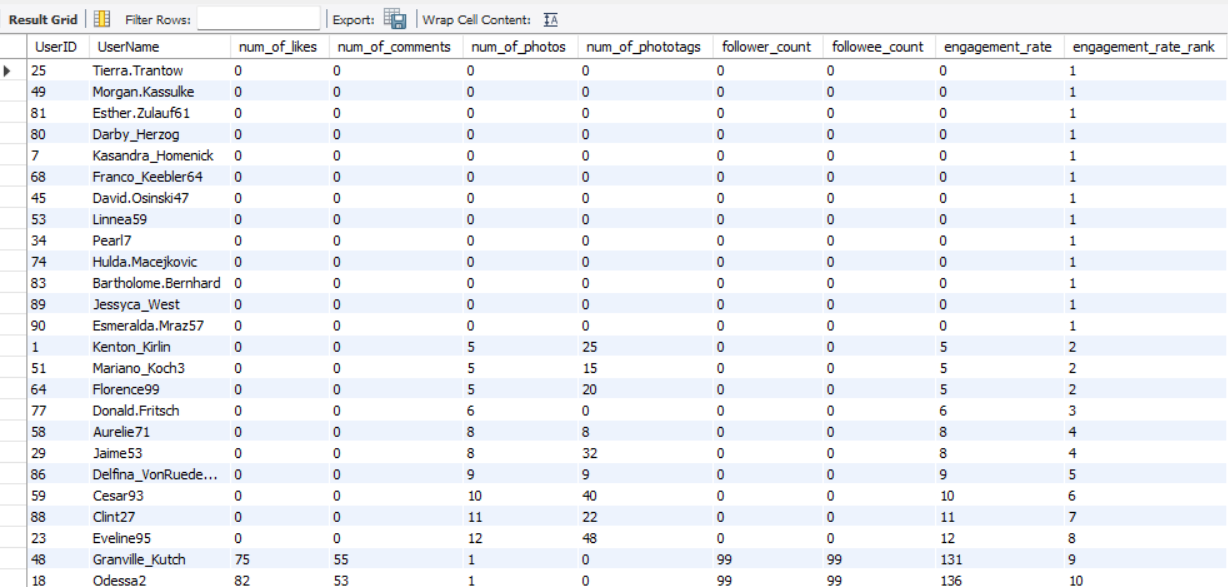
8. Calculated an engagement\_rate as the sum of likes, comments, photos .

9.Then ranked users by engagement\_rate in ascending order using **DENSE\_RANK().**

10.This ranking will give us the inactive users.

**This is how we will rank users based on user engagement and activity levels. The highly rank users can be considered as inactive users.**

****

**Result –**

**Strategies recommendation for re-engaging users:**

**1.Re-engagement Campaigns:** These campaigns may consist of time-limited deals, special discounts, or freebies that require interaction, such as commenting, sharing, or tagging a friend.

**2.Re-engagement E-mails:** Send personalized emails showcasing trending content or posts from their connections that they’ve missed. Suggest content, groups, or hashtags they might be interested in based on their past behaviour.

**3.By Giving Discounts and Gift Card :** This is one of the most popular yet effective ways to re-engage user. Because they find these rewards exciting, this will encourage them to utilize the site more.

**4.Taking Feedback User Surveys:** Ask inactive users for feedback on why they stopped engaging. Use their input to make meaningful changes or offer solutions directly.

**5. Simplifying The Use :** Offer templates or prompts that make it easier for users to create and share content, reducing the friction of coming up with something new.

**6. Contests For Users** : Introduce challenges or contests with rewards for the most engaging content or for hitting specific activity milestones.

**7. Data-Driven Insights:** Use the old data of inactive users when they created account like whom did they follow first , which post they like first and so on.

These insights can help refine content strategy based on user preferences.

**These approaches can help re-establish relationships with inactive users, making them feel valued and more likely to re-engage again.**

**Task-3. Which hashtags or content topics have the highest engagement rates? How can this information guide content strategy and ad campaigns?**

**Ans –**

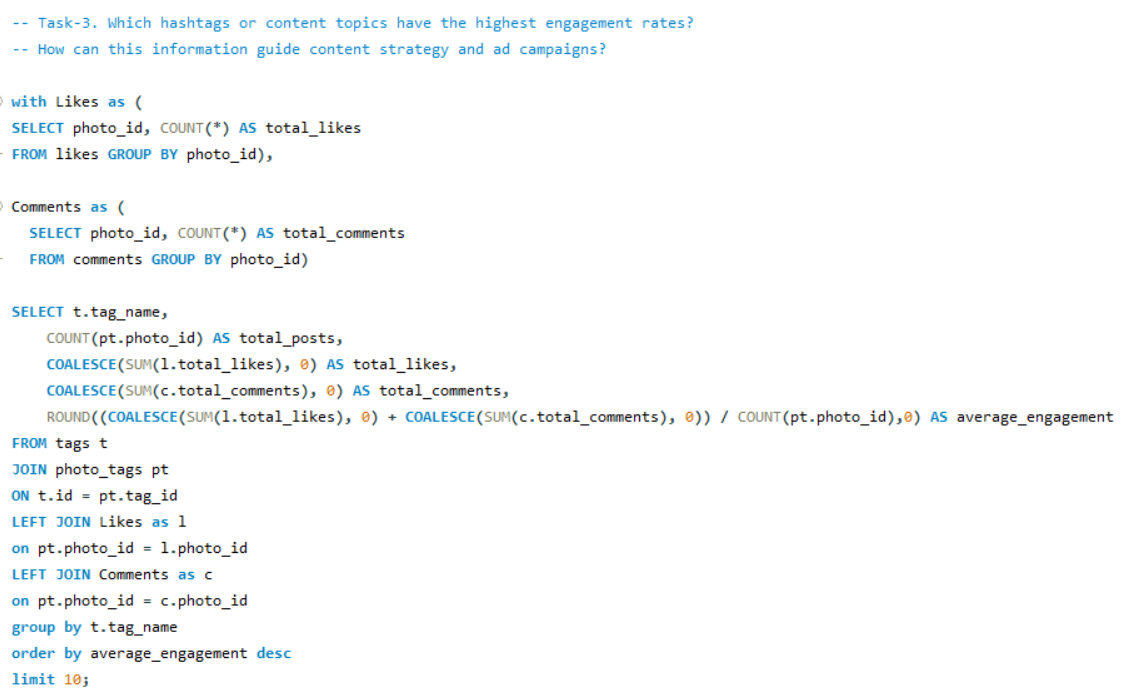
1.Firstly I created two CTEs **Likes** and **Comments** to count the total likes and comments per photo\_id.

2. I **join** tags table with photo\_tags to get each tag associated with a photo and then left join it with **Likes** and **Comments CTE’**s to bring in the total likes and comments for each photo.

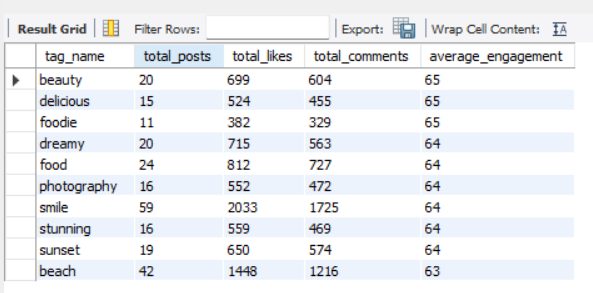
3. For each tag, I calculated the total number of posts, total likes, and total comments.

4. The engagement rate is calculated as the **average sum** of likes and comments per post for **each tag**.

5. At last, the results are ordered by engagement in descending order, with the **top 10 tags** displayed. This helps in identifying the most engaging hashtags to guide content strategy and ad campaigns.



**Result –**



**The Hashtags with the highest engagement rate are:**

1.beauty

2. delicious

3.foodie

4.dreamy

5.food

6.photography

7.smile

8.stunning

9.sunset

10.beach

**Strategies –**

**1.Prioritize topics :** Creating more posts around popular hashtags or high-engagement topics can help attract more followers and keep existing ones engaged.

**2. Targeted Ad-Campaign :** Audience is more likely to connect with ads that are related to interesting and well-liked subjects, which will boost ad performance and conversion rates.

**3. Refine hashtag’s :** One can boost organic reach by using insights about the most effective hashtags. With these hashtags it will increase post visibility and engagement.

**4. Tailored Recommendations:** Use insights from these hashtags to tailor content recommendations and suggestions for users, enhancing their overall experience and increasing engagement.

**5. Trending Content Series:** Develop a content series based on these popular hashtags. Example, a weekly #sunset or #party post could keep followers engaged and looking forward to new content.

**6. Cross-Promotion:** Cross-promote content , that features these hashtags across different social media platforms, driving more traffic and engagement back to Instagram.

**7.Influencer Collaborations**: Partner with influencers who frequently use these hashtags and have high engagement rates. For example, a collaboration with a food influencer under #delicious or #foodie can drive targeted traffic and engagement.

**8. Audience Insights :** Topics with high levels of engagement show what appeals to audience most and provide hints about their interests, values, or current trends. More individualized material that fosters loyalty is made possible by this understanding.

**Task-4. 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 –**

1.The query calculate **posts, likes, and comments** over time to determine photo engagement.

2. In order to determine the overall number of likes and comments for every photo, it first generates temporary tables CTE’s (**Likes and Comments**).

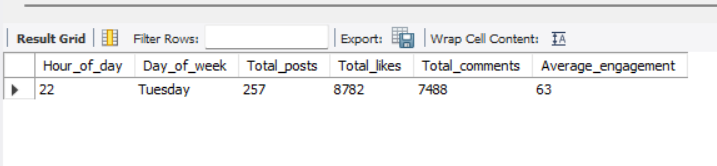
3. The primary query then extracts the **hour** and **day** for grouping by joining this interaction data with the publishing time of each photo.

4.We can observe when engagement is at its peak because it **groups** the data by hour and day of the week.

5.After calculating the total number of posts, likes, and comments for each time period, it divides the **total number of likes and comments** by the total number of posts to determine the **average engagement**.

****

**Result -**

****

**Pattern -**

1.The above output shows that most of the users are engaged at 10 PM on Tuesdays with an average engagement of 63 .

2.With a total of **257 posts**, **8782 likes** and **7488 comments** the average **engagement rate** is **63** for nearly 100 users who majorly post at **10 PM** on **Tuesdays** is calculated.

3.Therefore according to the analyses since the user engagement is high mostly during that day of week and hour this pattern can be used for targeted ad marketing.

**Strategy for target marketing –**

1.There are definitely patterns and trends in user engagement based on the demographics.

**2. Active Topics and Hashtags :**The night posts, reels, videos use hashtags like #party\_night, #moon, #dinner #soothing\_songs ,#party , #concerts etc, this content will gain more reach when posted in the night than on the other time of the day since it will be more relatable at that time of the day.

1. **3.Targeting Different Age Group :**Similarly age and gender also matter because mostly different age groups attract different and seek different contents. The adults will gain more content of enjoyment (party, late night drives , movies, concerts, etc) , while aged people will like to see old songs reels , bhakti posts and ladies will consume content of styling , makeup , cooking content and etc.
2. **4. Location Targeted Marketing :** Also location plays an important role For example, when a content is related to Japan various hashtags come under its umbrella like #anime,#sakura,#Tokyo,#lifeinJapan,#buddha,#ramen,#sushi etc. These tags are related to the location.
3. **5.Content Personalization :** Create content that resonates with specific demographics based on their engagement patterns. Tailor messaging and imagery to match the preferences of different groups.
4. **These insights will help to reach the targeted audiences easily.**

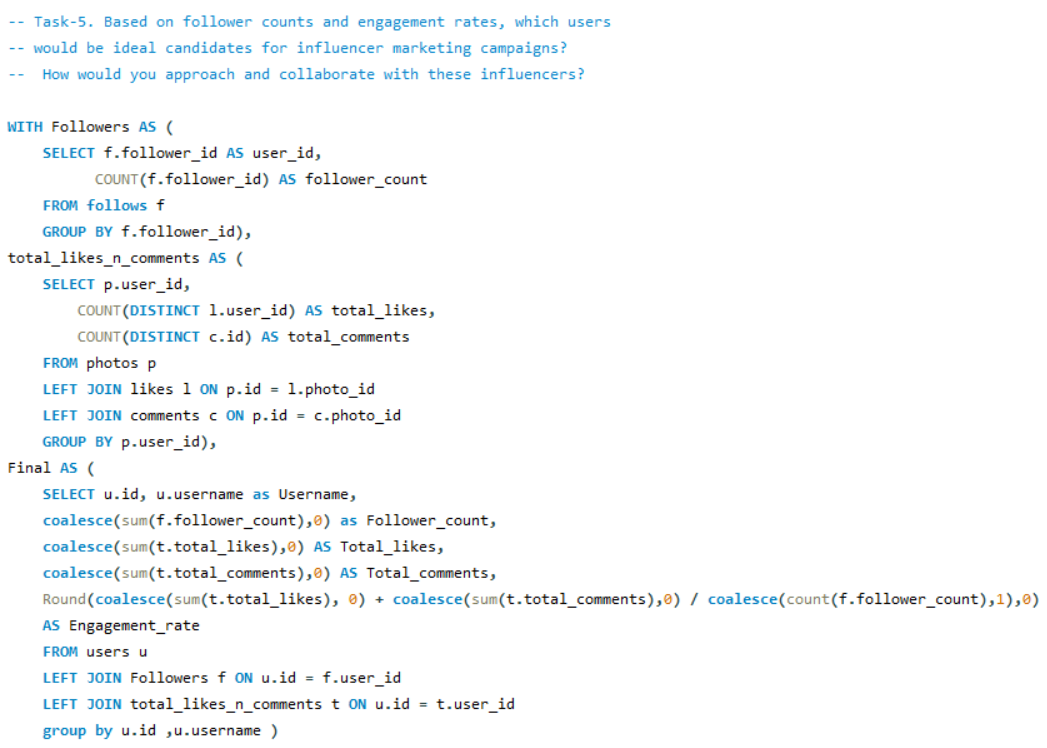
**Task-5.** **Based on follower counts and engagement rates, which users would be ideal candidates for influencer marketing campaigns? How would you approach and collaborate with these influencers?**

**Ans –**

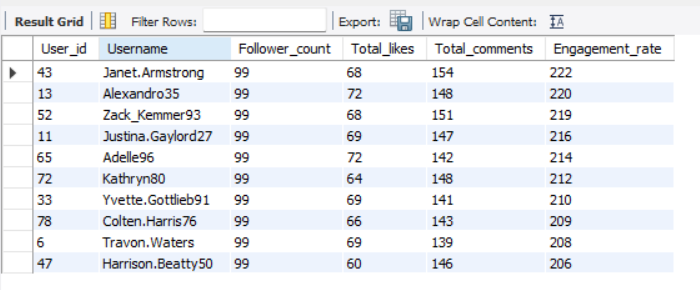
1.Through an analysis of follower counts and engagement rates, this query determines the **top 10 users** for influencer marketing.

2. In the Followers temporary table, it first determines how many followers each user has, and in the **total\_likes\_n\_comments** table, it totals the number of likes and comments for each user.

3. This data is then combined in the final table, which uses **COALESCE** to handle any missing values and divides the **total number of likes** and **comments** by the number of followers to get each user's engagement rate.

****4.The top ten possible influencers for cooperation are returned after filtering for persons with followers and **sorting** them by **follower count** and **engagement rate** (highest first).

****

**Result -**

**Graph for Follower Count and Engagement Rate -**

**Approach and collaboration –**

**1.Identify Potential Influencers**: Look for users with high follower counts and strong engagement rates. Understand the influencer’s content and audience. Prioritize those whose content aligns with your brand values and target audience.

**2.Collaboration Strategy:** After selecting the influencers approach them with clear objectives for the campaign. Offer opportunities like product collaborations, sponsored content, or exclusive access to events or features. Give them creative freedom within content guidelines to maintain authenticity.

**3.Build Relationship with Influencer :** Build long-term relationships with influencers rather than one-off campaigns for sustained engagement. Consider offering affiliate programs or ambassador roles.

**4.Compensation & Incentives:** Provide fair monetary compensation or product gifting.

**5.Tracking & Amplification:** Set clear KPIs(Key Performance Indicator) to measure success. Amplify influencer content across your brand’s channels for maximum reach.

**This is how I would you approach and collaborate with these influencers.**

**Task-6.** **Based on user behavior and engagement data, how would you segment the user base for targeted marketing campaigns or personalized recommendations?**

**Ans –**

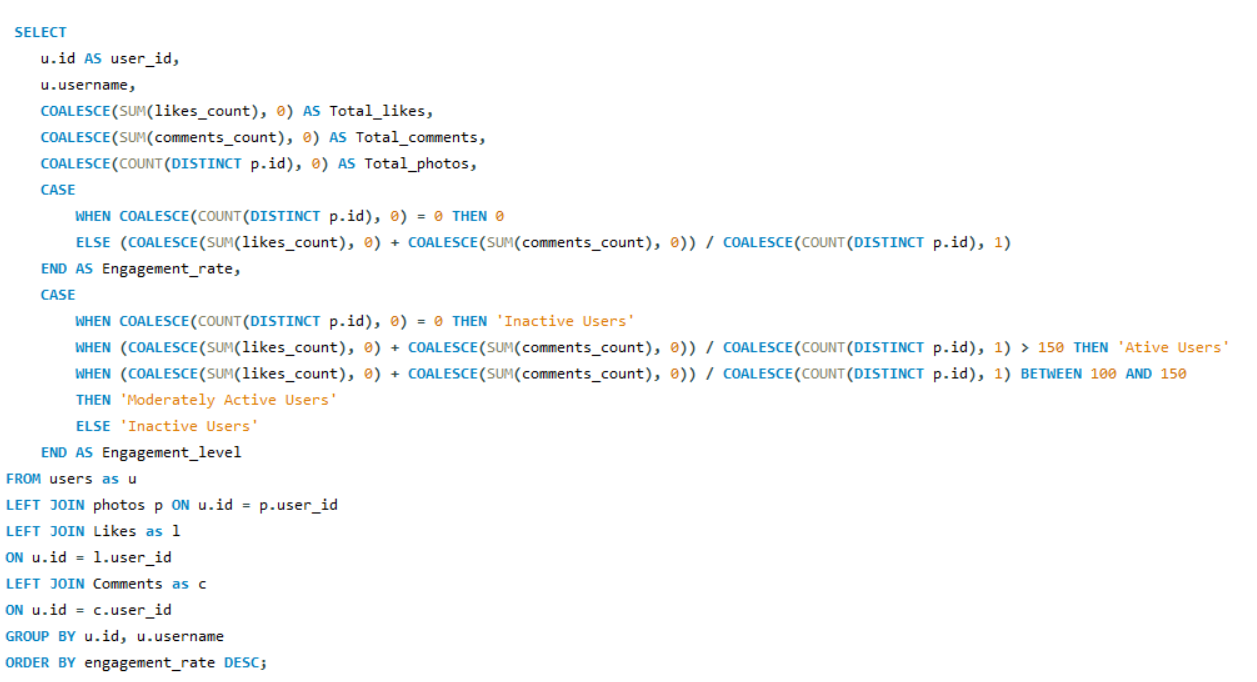
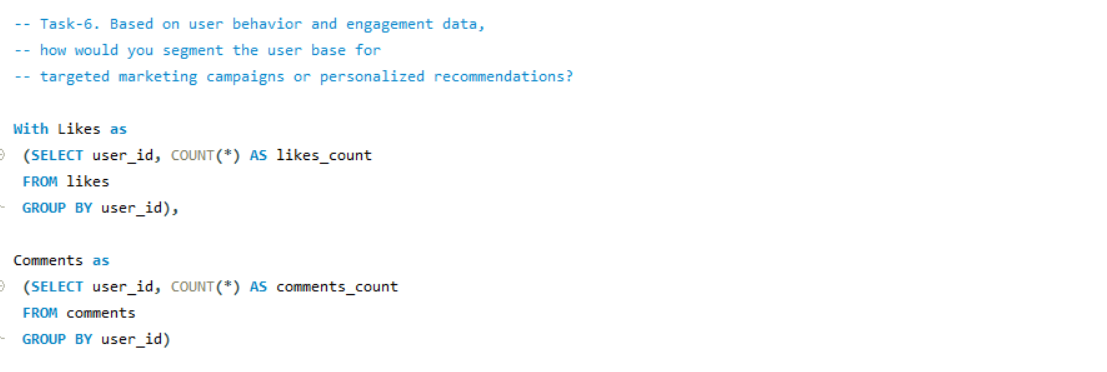
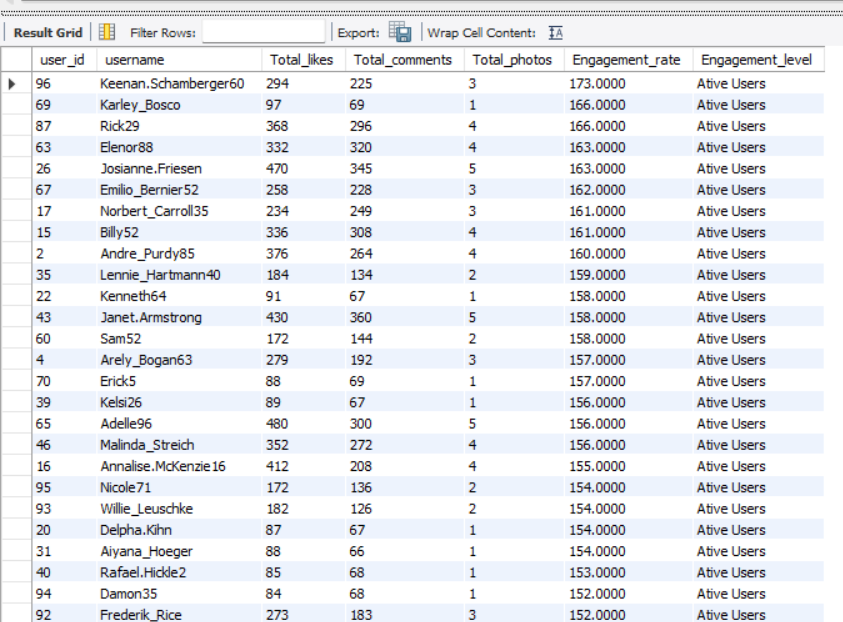
1.This query segments users based on the engagement levels of the users , categorizing them as **Active Users, Moderately Active Users, or Inactive Users.**

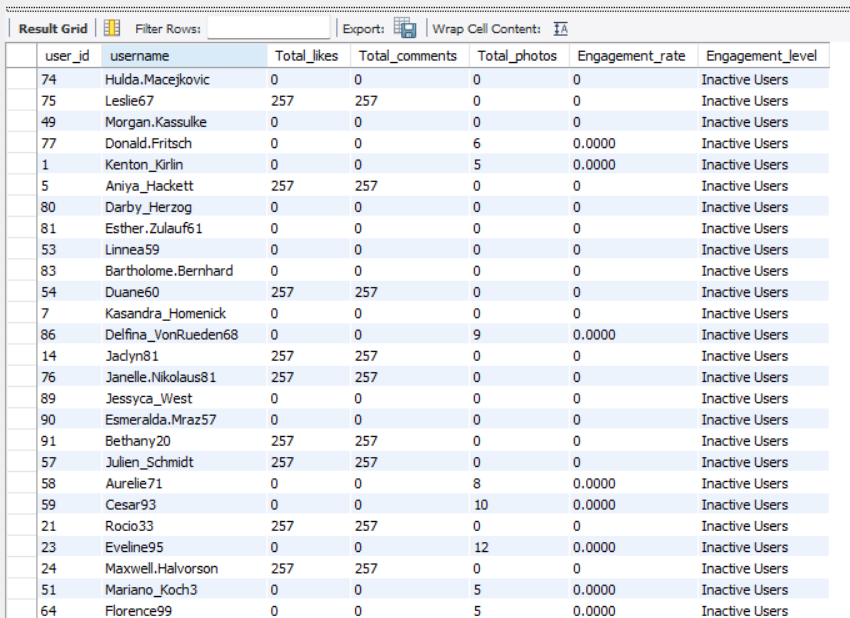
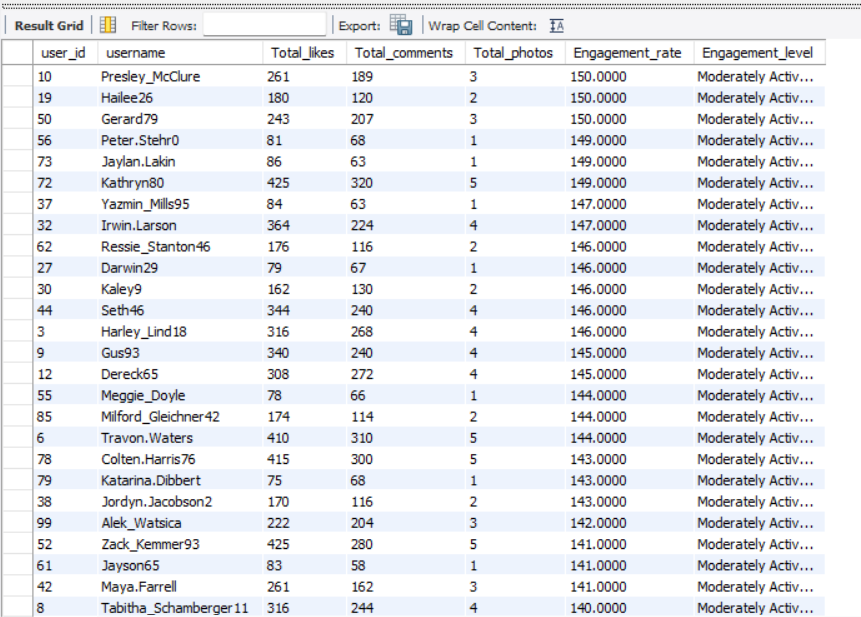
2.It first calculates the total likes and comments for each user using **CTE’s** temporary tables **Likes** and **Comments**. Then, it calculates the total number of photos posted by each user.

3. The Engagement\_rate is calculated by dividing the sum of likes and comments by the total number of photos.

4. Depending on this engagement rate, users are categorized into one of the three segments: users with an engagement rate **above 150** are considered **Active,** those **between 100 and 150** are **Moderately Active**, and **others** are marked as **Inactive**.

5.The query **groups** users by their **ID** and **username**, ordering the results by engagement rate in **descending order** to prioritize the most engaged users.

**Result -**



**Graph for Engagement Level -**

30% of the users are Active Users i.e, 30 out of total 100 users are highly engaged.

34% of the users are Moderately Active Users i.e, 36 out of total 100 users are medium- engaged.

36% of the users are Inactive Users i.e, 36 out of total 100 users are low- engaged.

**Targeted marketing campaigns or personalized recommendations based on engagement level:**

**A] Active Users :**

**Marketing Campaigns :**  Provide exclusive discounts or first use on new items. Emphasize high-quality features or goods that suit their hobbies and activity pattern. Provide fair monetary compensation or product gifting.

**Personalized Recommendations:** Make recommendations for goods, services, or content that is comparable to what they regularly interact with. Adapt updates and notifications to their past interactions and preferences.

**B] Moderately Active Users:**

**Marketing Campaigns :**  To promote greater interaction, use targeted advertisements for interactive content or exclusive discounts. Offer them incentives, such as temporary discounts, to boost their level of interaction.

**Personalized Recommendations:** In order to encourage greater involvement, provide products or content that complement their previous behavior while also incorporating some fresh or popular choices.

**C] Inactive Users :**

**Marketing Campaigns :** Pay attention to re-engagement tactics. To regain their interest again, provide discounts, reminders, or informative materials. Think of employing win-back efforts with customized messaging.

**Personalized Recommendations:** Introduce them to well-liked or captivating products or information. To get their attention and motivate them to engage more, use prompts or well chosen content.

**Implementation Steps:**

**1.Analyze User Data :** Analyze user activity and engagement metrics on a regular basis to spot trends, improve segmentation, and adjust to changing patterns.

**2. Create Targeted Messaging :** Create customized messages and offers that speak to the particular tastes and traits of each user group.

**3. Conduct Strategy Testing :** Within each part, try out different tactics to see which ones have the most effects.

**4. Track and Optimize :** Keep an eye on the effectiveness of recommendations and marketing performance at all times. To improve results, adjust strategy based on comments and insights.

**This is how we can create segment of users and by personalized campaigning and recommendations we can build stronger customer relationships and long-term loyalty.**

**Task-7.** **If data on ad campaigns (impressions, clicks, conversions) is available, how would you measure their effectiveness and optimize future campaigns?**

**Ans-**

Analyzing a variety of indicators that reveal how well an ad campaign is reaching and interacting with its target audience is crucial for determining its efficacy. In addition to assessing the campaign's effectiveness, the goal is to use the data to improve  future campaigns for better results.

**Measuring Campaign Effectiveness:**

1. **1.Impressions**: Good visibility is indicated by high impressions. However, effectiveness is not fully revealed by impressions alone. They must be examined alongside other metrics, such as conversions and clicks.
2. **2.Clicks**: A greater click-through rate indicates that the targeting, creativity, and ad content are appealing and relevant to the target group of people. A low rate of clicks could mean that the wrong people are seeing the advertisement or that it is not engaging.
3. **3.Conversion Rate**: A high conversion rate shows that people are motivated to take further action in addition to being interested enough to click on the advertisement. A low conversion rate can indicate problems with the offer, the landing page, or the user experience in general.
4. **4.Engagement Metrics**: Even if it doesn't result in conversions right away, high engagement shows that the audience finds the ad content compelling, which can raise brand awareness and loyalty.

**Optimizing Future Campaigns:**

1. **1.Ad Testing:** Test multiple versions of your advertisement to determine which works best (e.g., headlines, pictures, calls to action). Make future ads better by using the results. Put more effort into the ad variant that has a greater click-through rate or conversion rate.

**2. Ad optimization :** Regularly refresh ad creatives (images, videos, copy) to prevent ad fatigue and maintain user interest. Use insights from top-performing creatives to inform design and messaging for future campaigns.

**3.Refining Targeting:** Examine which audience segments ( age, gender, geography, and interests ) are reacting most favorably to your advertisements. Targeting should be modified to concentrate on high-performing segments. Spend a larger portion of your advertising budget on targeting a specific demographic if statistics indicates that they convert more frequently.   
  
**4.Landing Page Optimization:** Continuously improve landing pages based on user behavior data (e.g., heatmaps, session recordings). Align landing page content with the ad's message to increase relevance and conversion rates.

**5.Utilizing Retargeting:** Show advertisements to people who have already interacted with your brand but haven't converted by using retargeting. Retarget those people with tailored advertisements to convince them to finish the conversion if data indicates a significant decrease following the initial engagement.

1. **6.Continuous Monitoring and Iteration:** Keep a close eye on the effectiveness of  campaigns and be prepared to make changes as necessary. Don't be afraid to make real-time adjustments to your budget, targeting, or advertisements if you observe a drop in performance indicators.
2. **7.Budget Allocation:** Make budget for the campaign and spend accordingly on advertisement , for collaborating with influencers , etc. Also make changes in budget if necessary to reach maximum people and make campaign successful .

**We can make data-driven decisions to optimize future ads and guarantee more engagement, conversions, and overall return on investment by methodically evaluating the success of our ad campaigns using these metrics and tactics.**

**Task-8.** **How can you use user activity data to identify potential brand ambassadors or advocates who could help promote Instagram's initiatives or events?**

**Ans –**

The users are can be divided into three categories i.e Highly Engaged, Moderately Engaged, and Less Engaged on the basis of engagement rate and followers count .The users who fall into the Highly Engaged will be given priority as possible brand ambassadors or advocates.

**Here's how we can use user activity data to our advantage:**

**1.Identify Highly Engaged Users:** Keep track on users who frequently like, comment, share, and interact with your stories, since this indicates a consistently high level of engagement. These users exhibit sincere interest in the brand and actively engage with the content. Find users who mention the brand in their posts and articles or who share the brand's material on a regular basis. Sharing shows that people are not only interacting but also contributing to the brand's message being heard more widely.

**2.Check Influence and Reach :** Seek out users who have a sizable following because they may be able to reach more people. Although having a lot of followers isn't enough, it's an excellent place to start. Determine if the user is already well-known in a related sector. Influencers can successfully promote projects or events if they share the brand's values.

**3.Check Long-term involvement and Loyalty** **:**  Find users who have been interacting with the brand regularly for a long time. Long-term involvement shows brand loyalty and a stronger bond with the company. Seek out users who have previously promoted the brand without receiving any compensation. They are excellent candidates for official ambassador positions because of their sincere support.

**4.Check Content Alignment with Brand Values:** Examine the kind of material the user frequently uploads. Are they producing material that is consistent with the values and image of the brand? Instagram users that regularly post about fitness, health, or mindfulness, for example, can be perfect ambassadors for a wellness event. Seek out those whose postings exhibit originality and ingenuity. When shared, excellent, eye-catching material can improve the brand's reputation.

**5.Personalized Outreach:** Contact them with a message that highlights their influence and how they align with Instagram's initiatives.

**6.Value Proposition:** Offer incentives such as exclusive access, features, or rewards for participating in campaigns.

**7.Collaboration:** Set Clear Expectations: Define the campaign goals, the content you expect them to create, and how they can share their involvement.

**8.Offer Creative Freedom:** Allow them to create content in their style to maintain authenticity.

**9.Long-term Partnership:** Consider establishing ongoing relationships with top-performing ambassadors to promote future initiatives or events.

**Task-9.** **How would you approach this problem, if the objective and subjective questions weren't given?**

**Ans –**

Well certainly the Objective and Subjective questions were very helpful as they provided the outline and better understanding of the problems and how to reach the solutions. But if the case was that the Objective and Subjective questions were not given then I would have followed a certain approach with certain steps to analyse and come up with insights from the given data.

**The Approach and the steps in it would be as follows:**

**1.Problem Identification:** I would start by understanding the goals and desired output of the analysis, whether we are optimizing user engagement, increasing retention or identifying potential influencers. After analyzing all this I would set up my plan and course of action on how I am going to manipulate the data.

**3.Understanding the Data:**  The data is important in analysing the requirements of the question asked. In this case the data is insufficient and has data of 100 users only

**4.Data Cleaning and validation**: I would start working on the data and my first step would be to check for null values and duplicates in all the tables and if found then rectify them. So that data can be ready to work on.

**3.Identifying Key Metrics:**

**Engagement metrics:** The engagement rate of each user by calculating the likes, comments and tags by each user and photos. Also analysing the followers and following count of each user to identify rising influencers and creators.

**Retention metrics:** Identifying active users who constantly interact with users constantly and keep them engaged with their posts over various periods of time(weekly,monthly,yearly). Also identifying inactive users will help create campaigns and events to keep them more active by understanding their needs and requirements.

**Acquisition metrics:** By analysing the growth of new users and identifying how the high engagement users influence the growth of new users.

**4.Trend Formation:** Develop hypotheses based on observed trends or known business goals. For example, "Users who engage more frequently are more likely to retain." Then check for their followers' activity and various other things from which we can come to conclusion whether to work with them or not.

**5.Analysing Trends and Their Influence on High User Engagement:** Identifying the trends by analysing the hashtags used by the users and also analysing various metrics of it i.e, most liked hashtags, most used hashtags. This will help us understand what are the topics that the users are interested in and consume more time on the platform. Using this data we can find influencers for each type of hashtags that are analysed and create a campaign for the targeted users.

**6. Data-Driven Solutions:** Then finally I would use SQL operations on the data to check whether my overview analysis stands correct or not and if not then what are the trends and how the current trends are different.

**7.Insights of Current Scenario:** I would take all the outputs from the SQL operations and create a report of current scenario and how much we need to work in which aspects. The analysis will give us the aspects in which we are lacking behind, So that the management can take proper actions to rectify them.

**7.User segmentation and targeted marketing:**

**Segment by Engagement Level:** Group users based on their activity (e.g., highly active, moderately active, and inactive users). Identify users with a high follower count and engagement rate for potential brand ambassador roles.

**Personalized Content:** Use segmentation to deliver personalized content recommendations to different user groups. Target inactive or low-engagement users with re-engagement campaigns, such as email reminders or in-app notifications.

**8. Making Report :** At last I will make a report of all analysis I done. In that I will show the method of action I followed to reach to this analysis. In that report I will mention the area where company need to work and will give suggestions for the retention of the inactive users as I discuss earlier in previous questions. And I will also make the dashboardof this data to make it visually present .

**This is how I will approach this problem, if the objective and subjective questions weren't given.**

**Task-10.** **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-**

**THE QUERY:**

UPDATE User\_Interactions

SET Engagement\_Type = 'Heart'

WHERE Engagement\_Type = 'Like';

**Steps To Make The Update:**

1.The **UPDATE** command will specifies the table which is being updated, here it is User\_Interactions.

2.The **SET** clause defines the new value for a column.

3.The **WHERE** clause will ensure that, where the Engagement\_Type is Like there only it will make changes.

4.This query effectively changes Engagement\_Type from "Like" to "Heart."