# **Objective Questions**

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

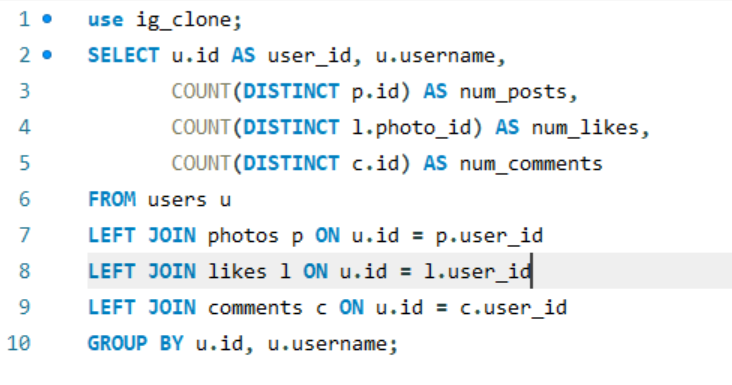
**Ans.**  I used the **GROUP BY** clause to group all the rows based on the specified column(s) — those that are being checked for duplicates or null values. The query results confirm that there are **No duplicate** entries or **NULL** values in those columns.

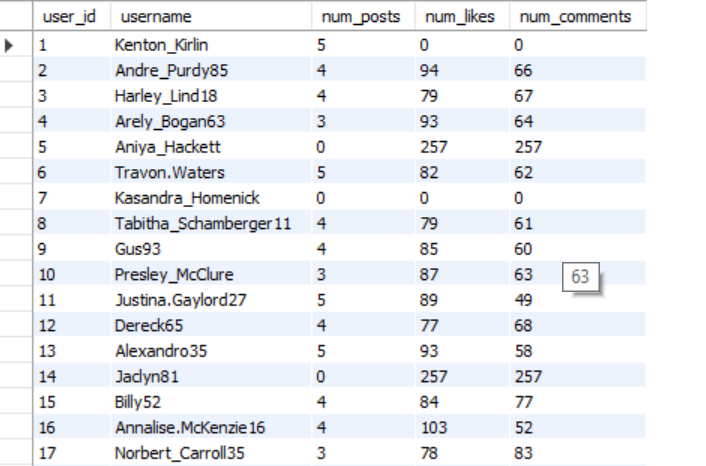


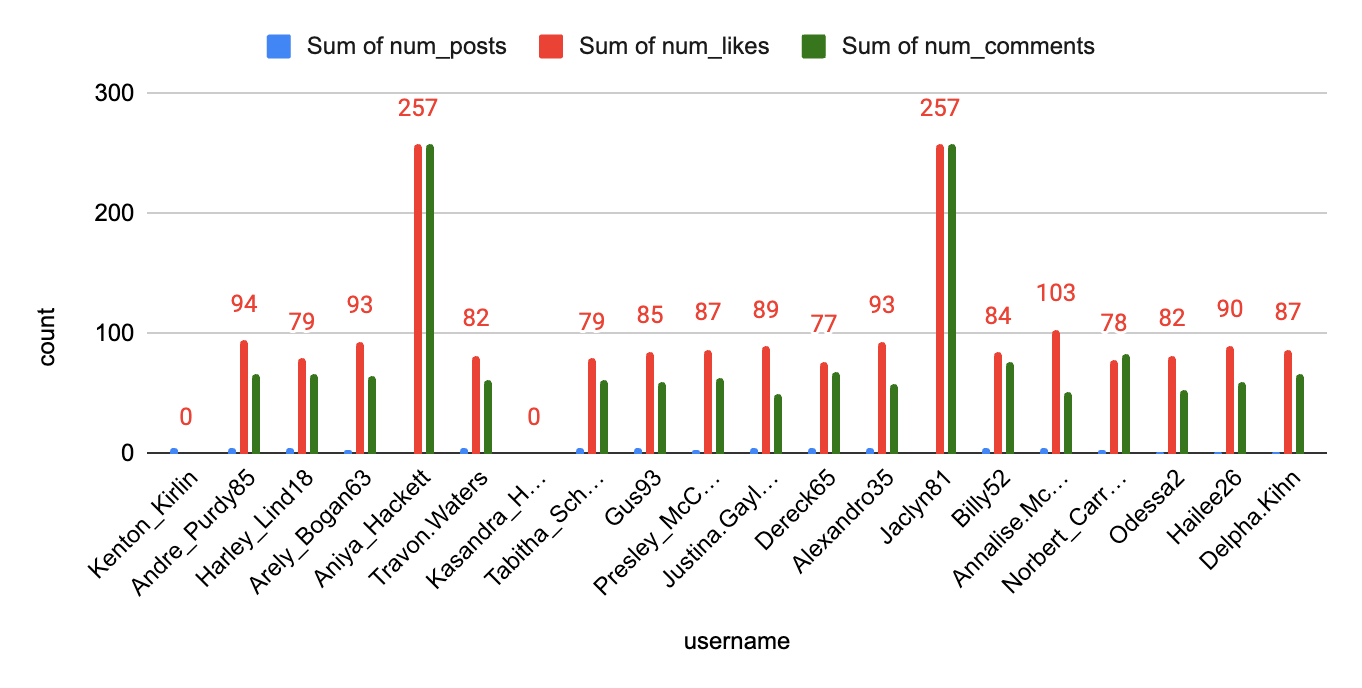
**Task-2.** What is the distribution of user activity levels (e.g., number of posts, likes, comments) across the user base?

**Ans**.

I wrote a query to calculate the total number of likes, comments, and posts made by users, aggregating data from the **photos, likes,** and **comments** tables.

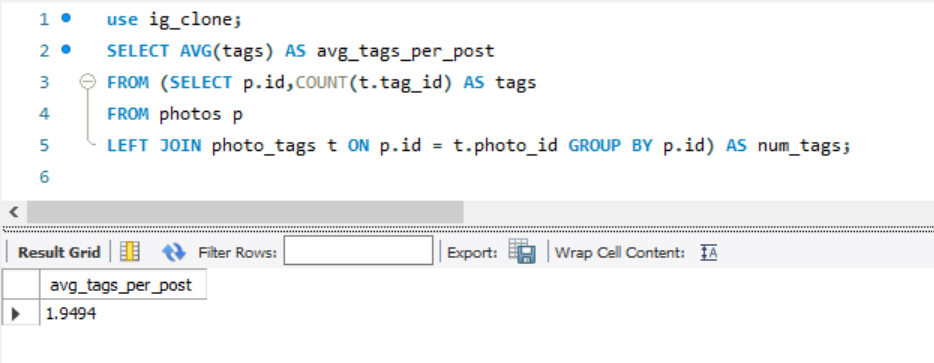


The distribution of user activity according to the likes, comments and number of posts as following:  




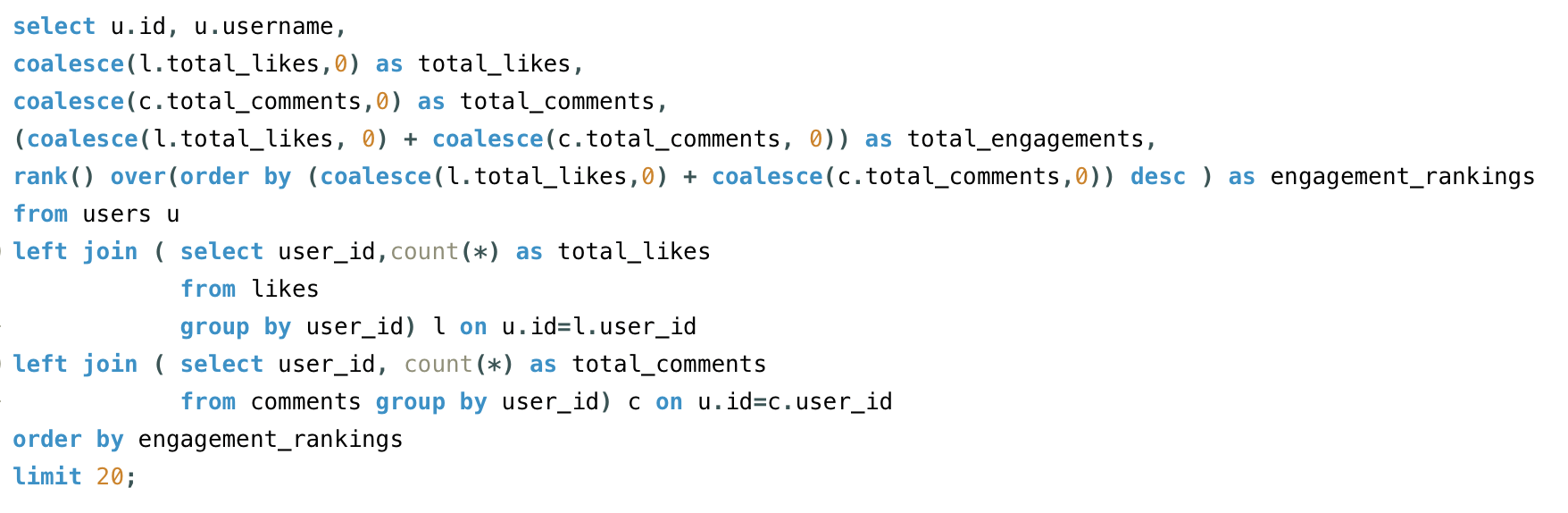
**Task-3.** Calculate the average number of tags per post (photo\_tags and photos tables).

**Ans.** I calculated the average number of tags per posts and the ans came **1.9494**.

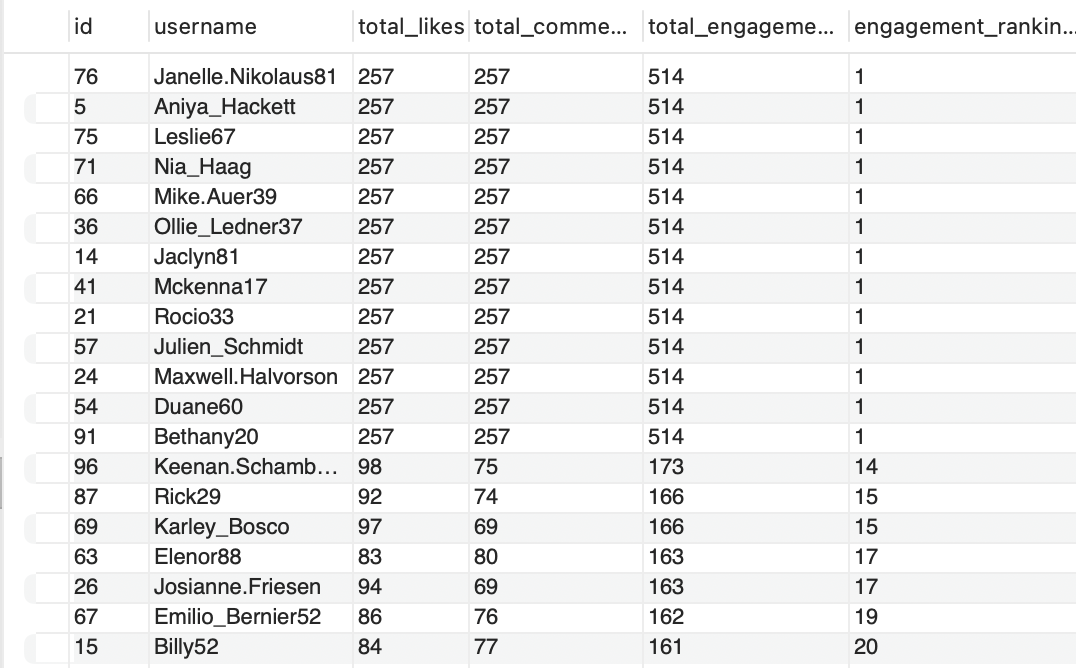


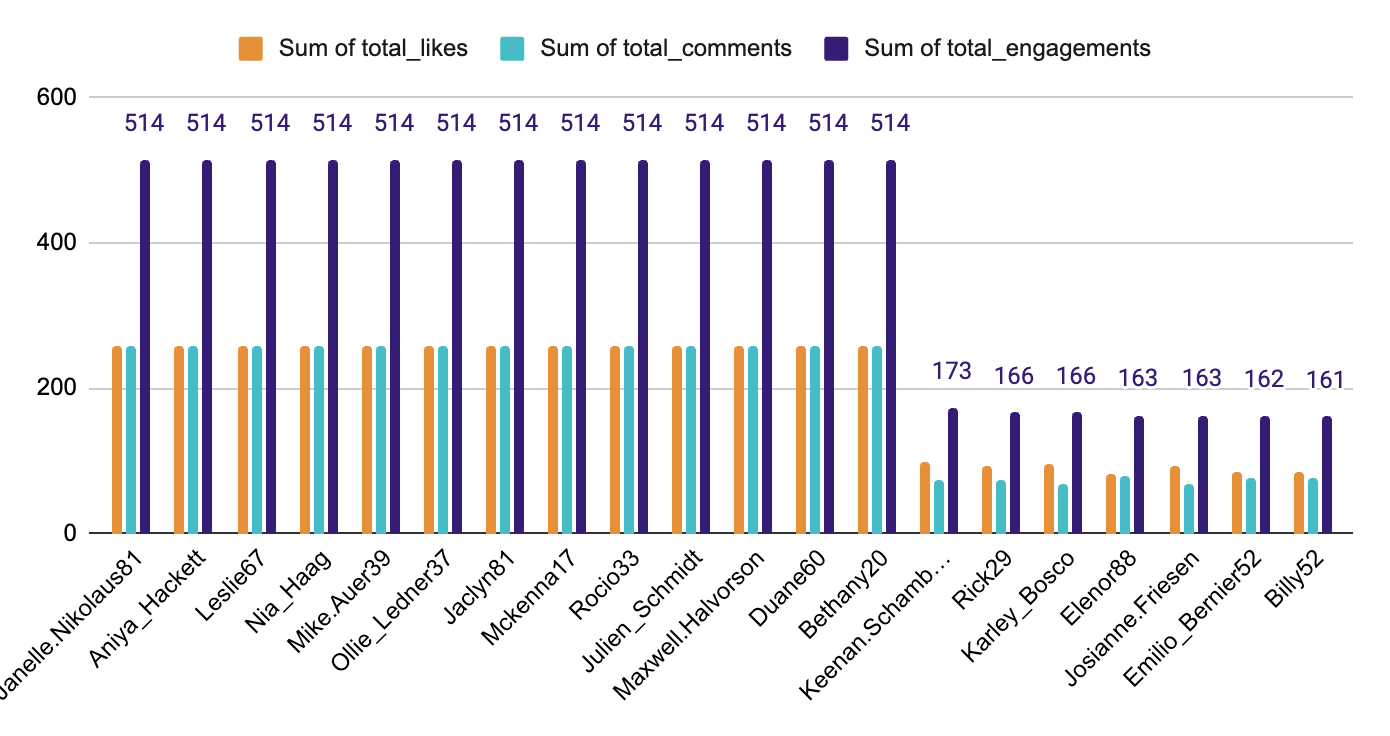
**Task-4.** Identify the top users with the highest engagement rates (likes, comments) on their posts and rank them.

**Ans.**

To identify the top users with the highest engagement rates, I wrote a query that creates temporary tables, l and c, through subqueries. In these subqueries, I calculated the total likes and total comments for each user\_id. I then joined these temporary tables with the main query to retrieve the user\_id, username, and their corresponding total likes and total comments.

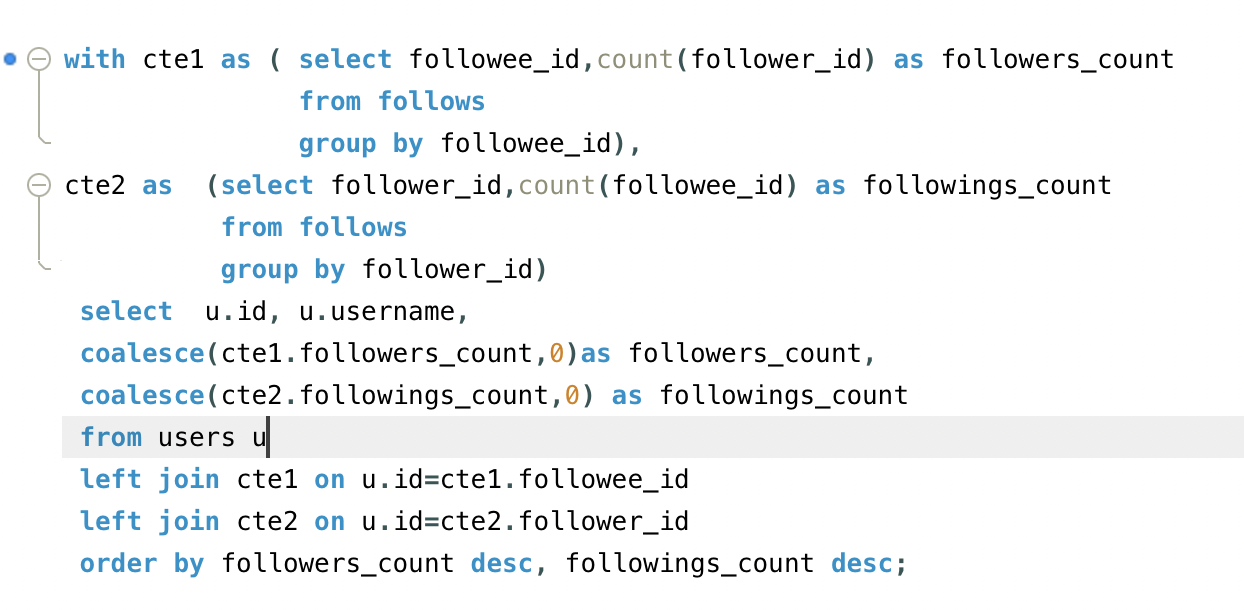
**Output:**

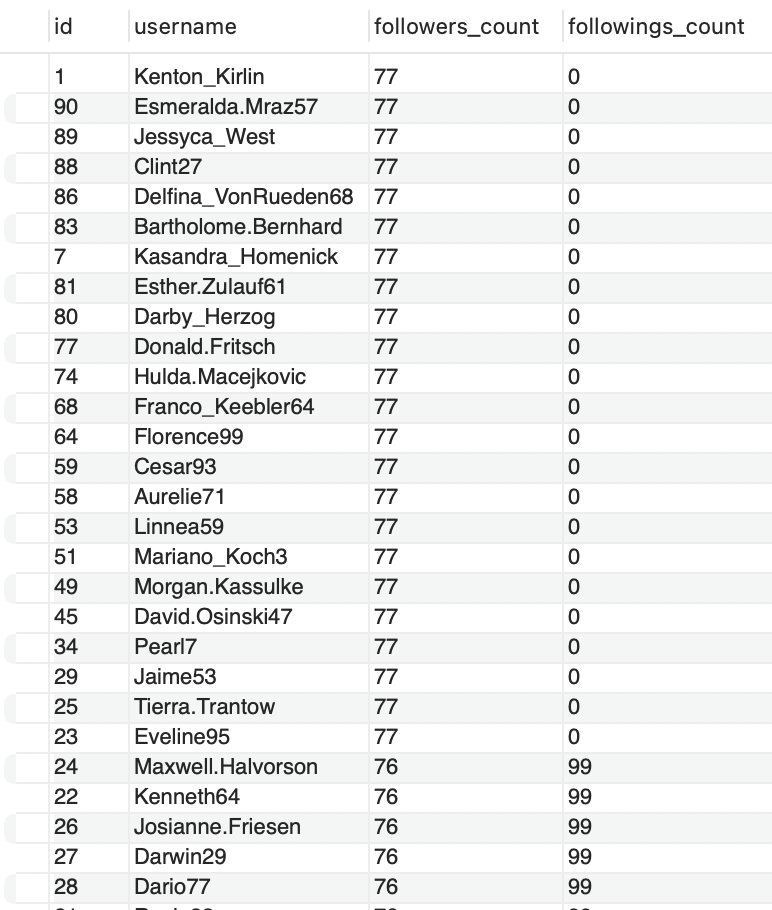




**Task-5.** Which users have the highest number of followers and followings?

**Ans.**

In this query, I utilized the concept of Common Table Expressions (CTEs) to construct temporary result sets that calculate the followee\_id along with their associated followers\_count, and the follower\_id along with their corresponding following\_count. Subsequently, I joined these CTEs with the primary query to extract the user\_id, username, followers\_count, and following\_count, thereby generating the resulting output.  
  


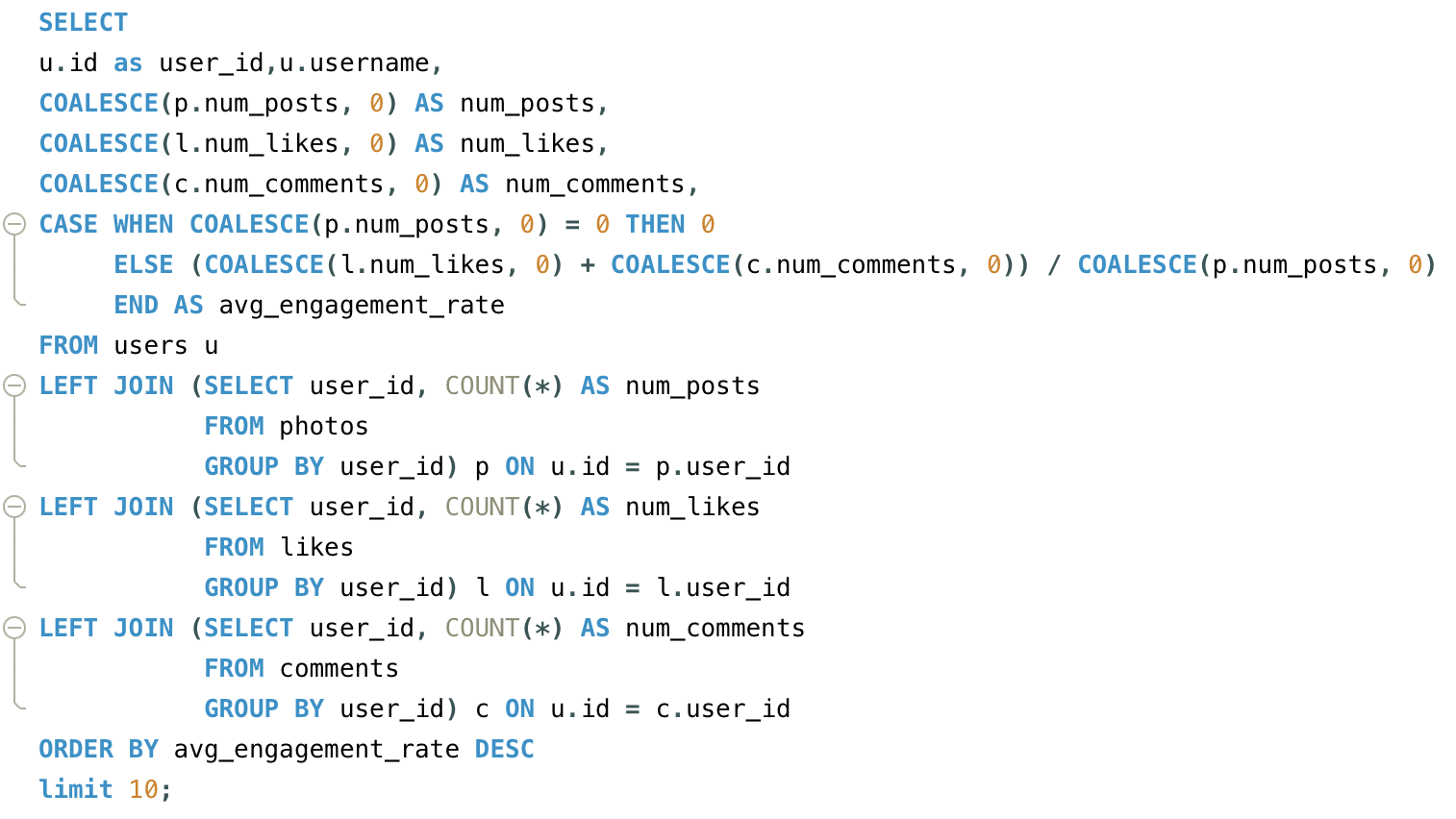


**Output:**

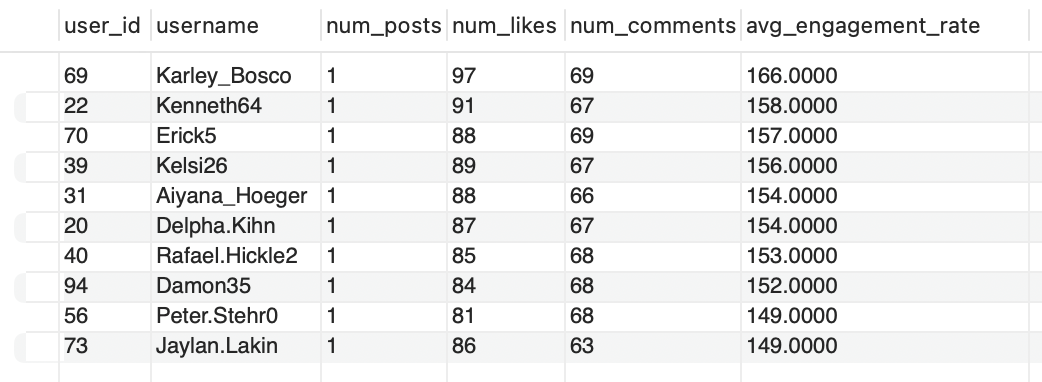
**Task-6.** Calculate the average engagement rate (likes, comments) per post for each user.

**Ans.**

I calculated the total number of posts, likes, and comments for each user using subqueries. These subqueries were then joined with the main query via left joins. Within the main query, I utilized the results of these calculations (total likes, comments, and posts) to compute the average engagement rate for each user.



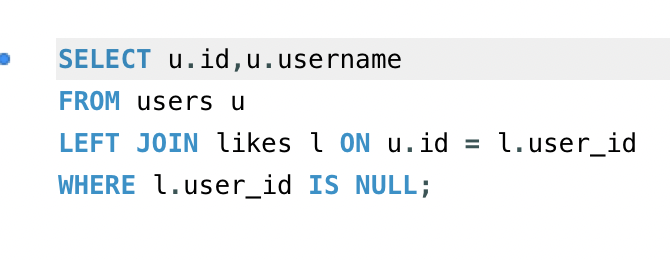
**Output:**

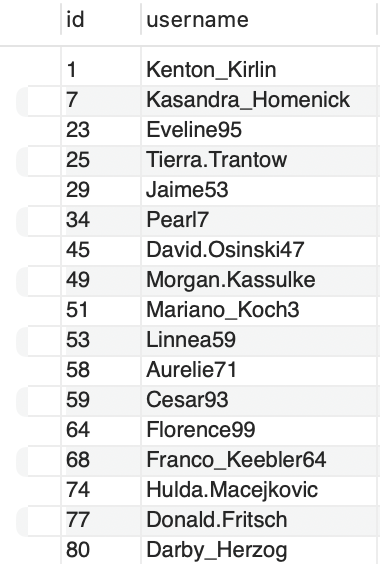




**Task-7.** Get the list of users who have never liked any post (users and likes tables)

**Ans.**

I calculated the id and usernames of those users who haven’t liked any posts , in which I joined the users table with the likes table. 



**Output:**

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

**Ans.**

User-generated content (UGC), including posts, hashtags, and photo tags, offers a rich source of insights that can be strategically harnessed to craft more personalized and engaging advertising campaigns. The following methodologies outline how this content can be leveraged to maximize user engagement and drive campaign effectiveness:

1. **Behavioral Analysis for Personalization:** By carefully analyzing user behavior—such as the posts they create, the hashtags they frequently use, and the content they engage with most—brands can design highly personalized advertising campaigns. This data-driven approach enables the delivery of relevant content that resonates with individual user preferences, fostering deeper engagement and increasing the likelihood of conversion.
2. **Identification of Trending Themes and Audience Sentiment:** Through the systematic examination of UGC, brands can uncover prevailing themes, trending hashtags, and popular content formats that are capturing the attention of their audience. These insights facilitate the creation of campaigns that align with the current interests and sentiments of users, thereby enhancing their relevance and emotional resonance. Such alignment amplifies the potential for user interaction and brand affinity.
3. **Enhanced Authenticity and Community Building:** Incorporating user-generated content directly into advertising materials not only enhances the authenticity of the campaign but also strengthens the sense of community between the brand and its audience. UGC serves as a form of social proof, showcasing real-world experiences and endorsements, which significantly boosts trust and relatability, making the brand appear more genuine and approachable.
4. **Tailored Communication via Notifications and Emails:** Leveraging UGC allows brands to send more tailored, contextually relevant notifications and emails to users. By utilizing insights from the content that users engage with most, brands can keep users informed about new developments, trending topics, or personalized offers aligned with their interests. This approach ensures that communication is not only timely but also directly relevant, increasing the chances of user engagement.
5. **Amplification Through Social Proof and Peer Influence:** User-generated content inherently carries the weight of social proof, as consumers tend to trust peer recommendations more than traditional advertisements. By integrating authentic user-generated posts, images, and testimonials into ad campaigns, brands can capitalize on this peer influence, making their messaging more compelling and credible.

**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.**

Correlation analysis can be effectively employed through statistical methodologies to assess the relationships between user engagement levels and the type of content consumed, such as photos, videos, or reels. If a robust correlation between user activity and specific content types is identified, brands can strategically prioritize content production that maximizes user engagement and satisfaction.

Upon conducting a thorough examination of the dataset and scrutinizing user behavior based on various content types (photos, videos, reels), along with hashtags associated with these formats, the analysis clearly reveals that certain genres of content consistently attract higher levels of user engagement. Notably, as engagement levels increase, there is a concurrent rise in the popularity of specific content types, indicating a direct correlation between user activity and content format.

This insight can be leveraged to refine content creation and curation strategies in several impactful ways:

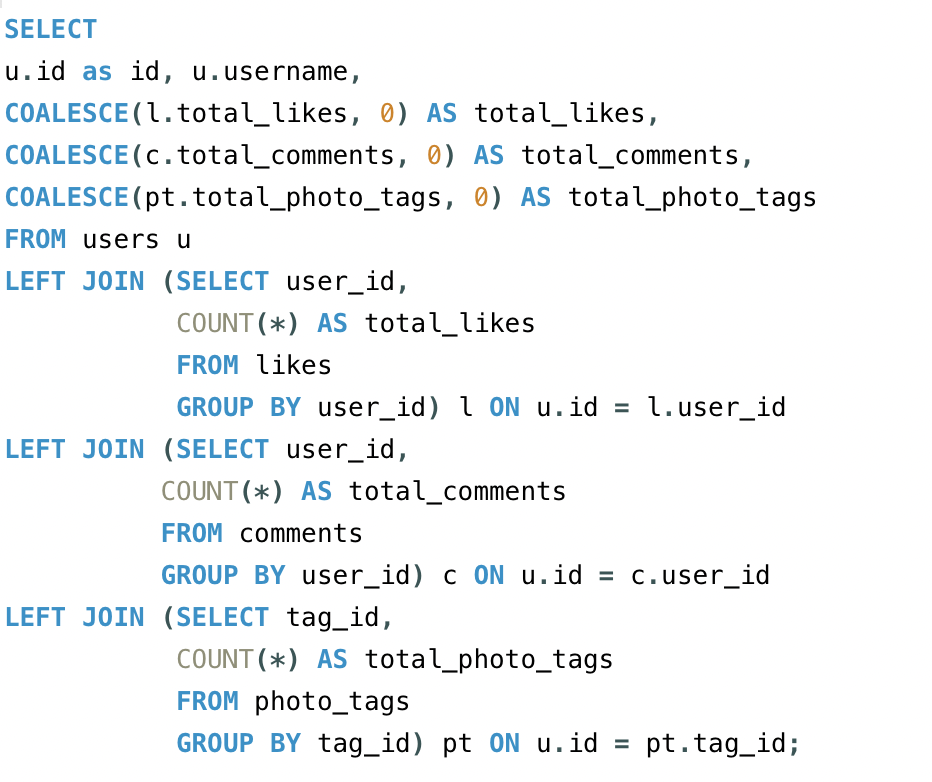
1. **Dynamic Hashtag Strategies**: To maintain relevance and capitalize on user interests, it is crucial to regularly update the hashtag list with trending and hot topics. This ensures that new and relevant content continues to capture user attention and aligns with current interests.
2. **Personalized Content Recommendations**: Leveraging data on user preferences and viewing history, brands can offer tailored content suggestions. By recommending content based on individual user interests—such as content they frequently engage with or follow—brands can foster a more personalized and immersive experience that drives further interaction.
3. **Utilizing Trending Hashtags**: Incorporating hashtags linked to emerging trends or hot topics within posts can significantly enhance engagement. By aligning content with widely discussed or popular hashtags, brands can increase their content's discoverability and reach, thus encouraging higher user participation.

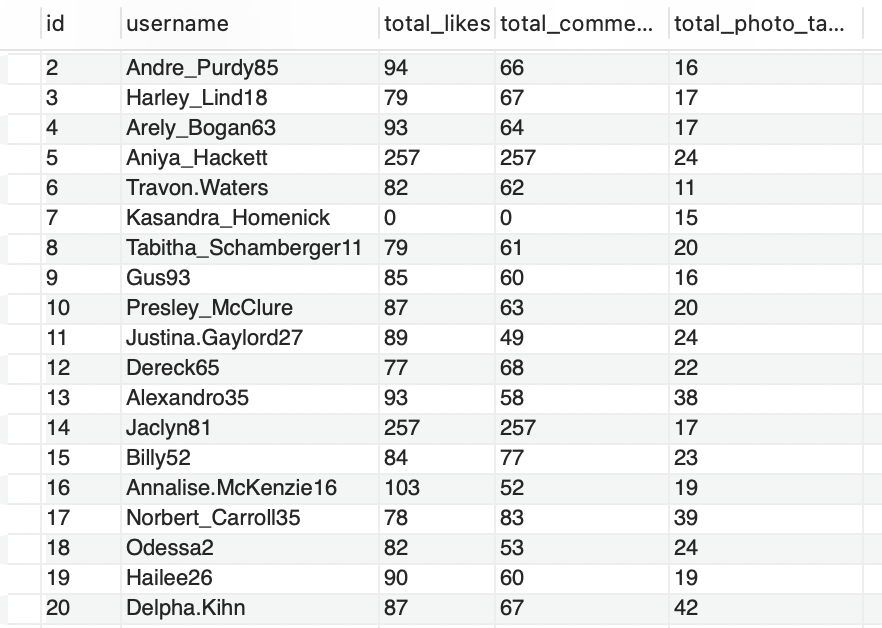
In summary, by analyzing user behavior and identifying content formats that generate the most engagement, brands can refine their content strategies. Regularly updating hashtags, personalizing content suggestions, and strategically using trending topics are all effective ways to optimize engagement and ensure content remains aligned with audience preferences.

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

**Ans.**

In this query, I employed subqueries and joins to calculate the total number of likes, comments, and photo tags for each user. These results were then integrated into the main query, which retrieves the user’s ID, username, and their corresponding totals for likes, comments, and photo tags.



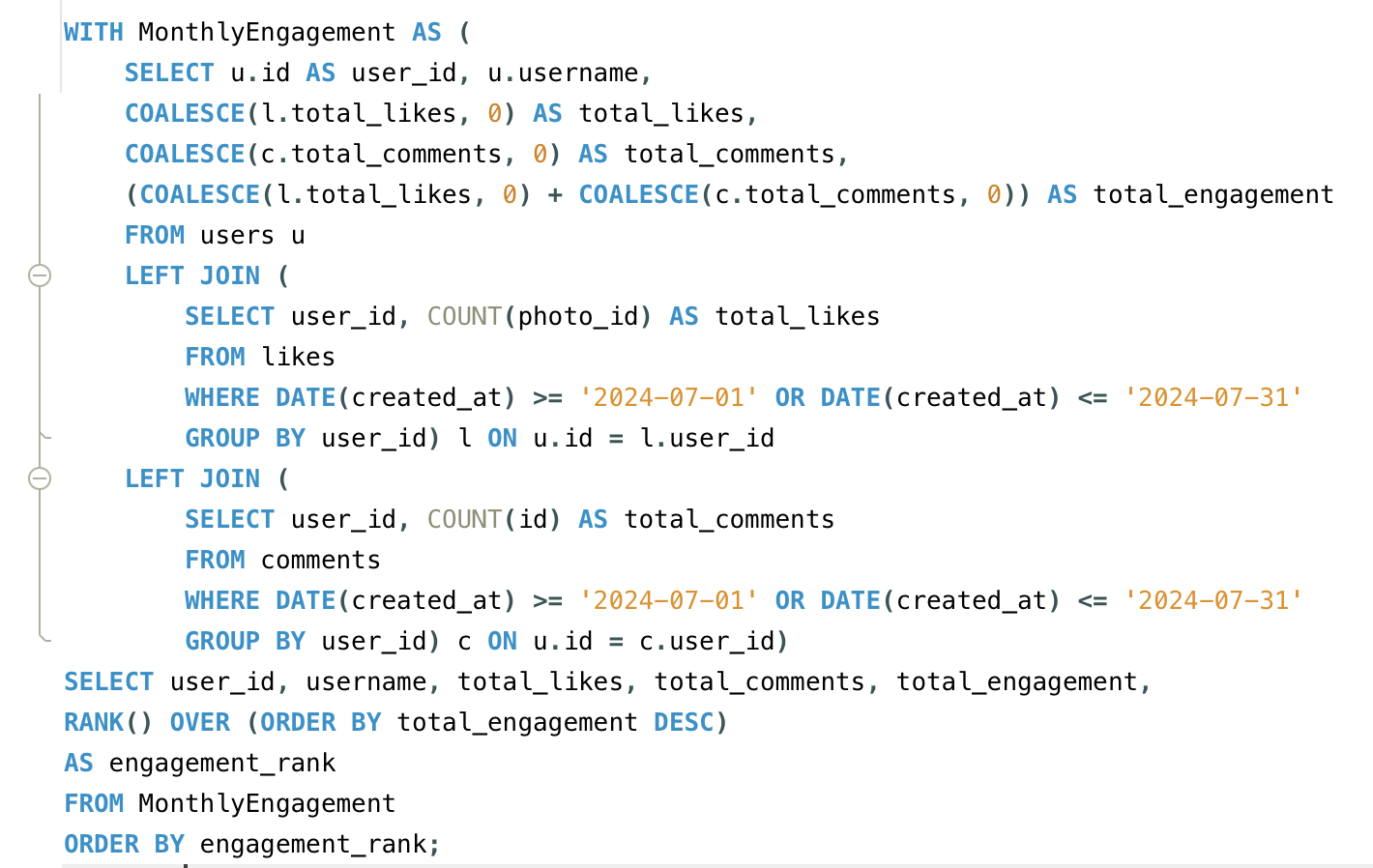


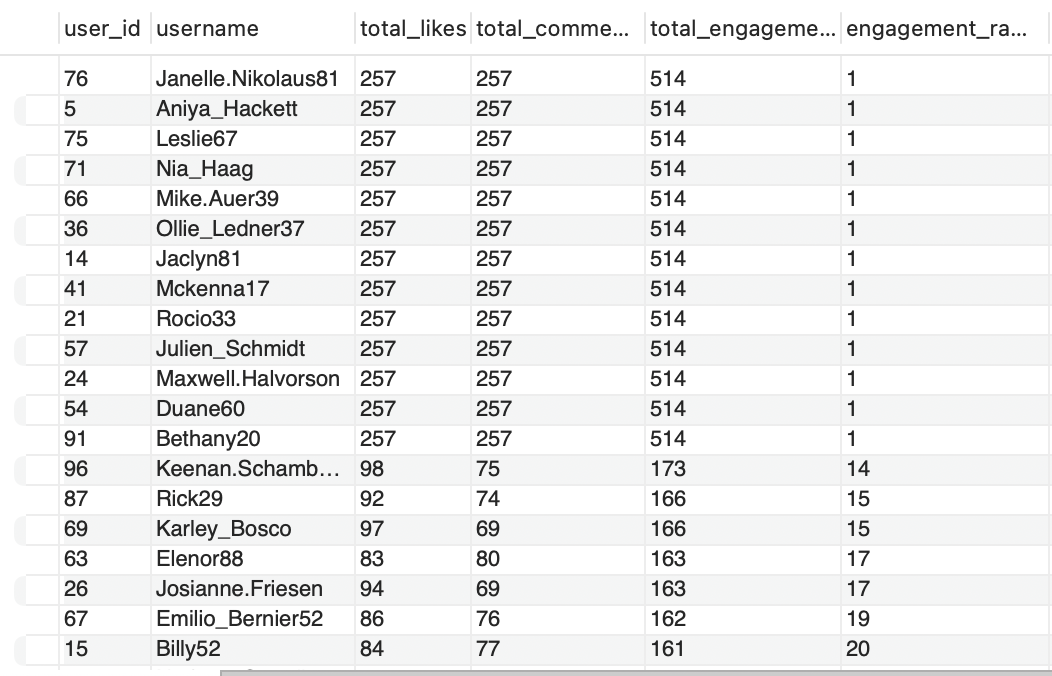
**Output:**

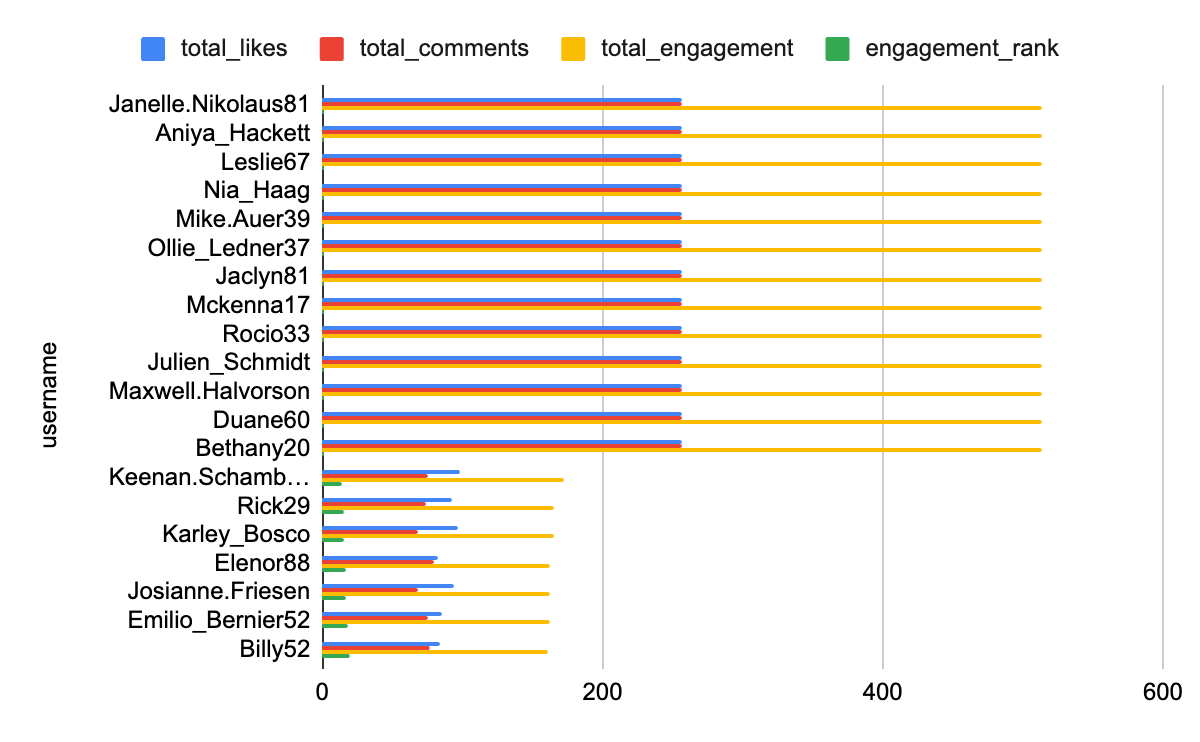


**Task-11.** Rank users based on their total engagement (likes, comments, shares) over a month.

**Ans.**

In this query, I applied a similar approach as before by utilizing Common Table Expressions (CTEs) to calculate the total likes and total comments. I then summed these values to determine the total engagement for the month of July. In the main query, I used the RANK() function to rank users based on their total engagement. The query first computes the total likes, total comments, and overall engagement, and then ranks the users according to their engagement levels.

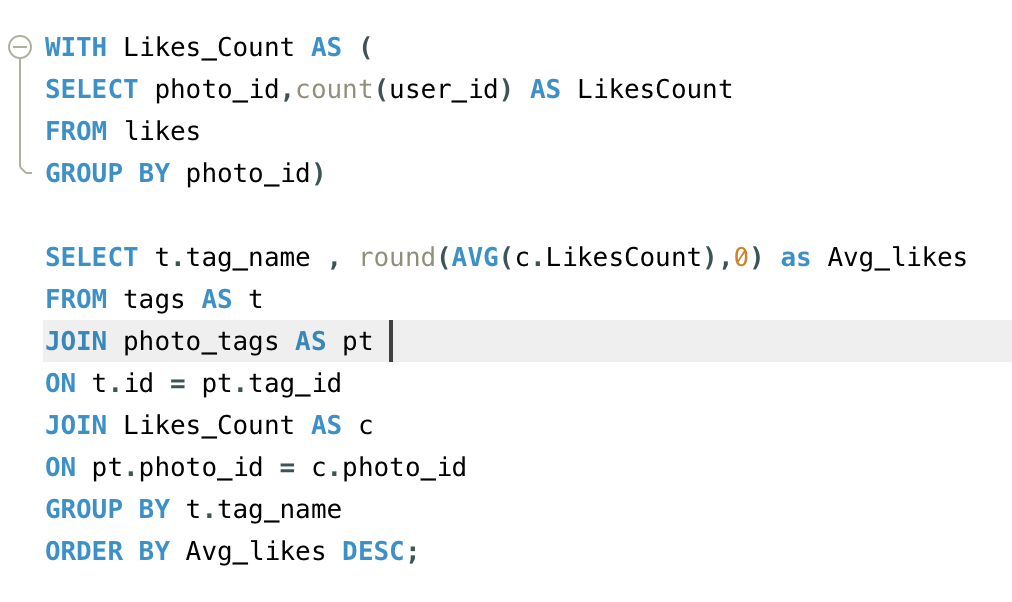
**Output:**

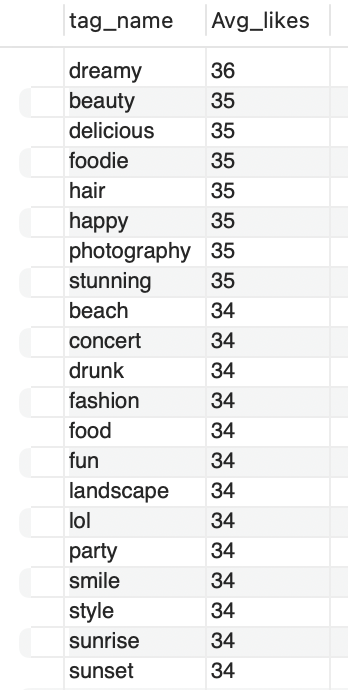


**Task-12.** Retrieve the hashtags that have been used in posts with the highest average number of likes. Use a CTE to calculate the average likes for each hashtag first.

**Ans.**

This query first calculates the total number of likes for each photo in the Likes\_Count CTE. It then joins the tags, photo\_tags, and Likes\_Count tables to link hashtags with the corresponding photos and their like counts. The query calculates the average number of likes for posts associated with each hashtag. It groups the results by hashtag name and orders them by the average number of likes in descending order. Finally, it returns the hashtags with the highest average likes.





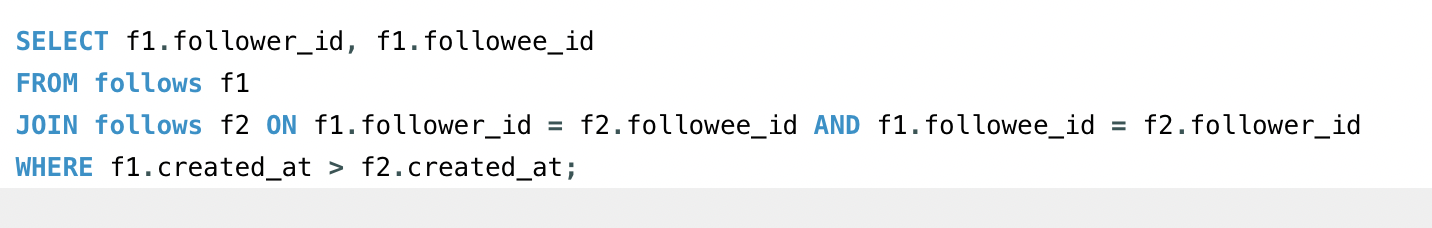
**Output:**

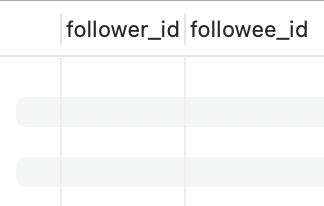


**Task-13.** Retrieve the users who have started following someone after being followed by that person

**Ans.**

To determine if there are any users who started following someone after being followed by that person, I employed a query that utilized the created\_at column to compare the timestamps of when each user was followed and when they started following others.

The query aimed to identify instances where a user was followed by another user before they themselves started following that same user. However, after executing the query, no such records were found, indicating that there are no users who started following someone after being followed by that person.  
  


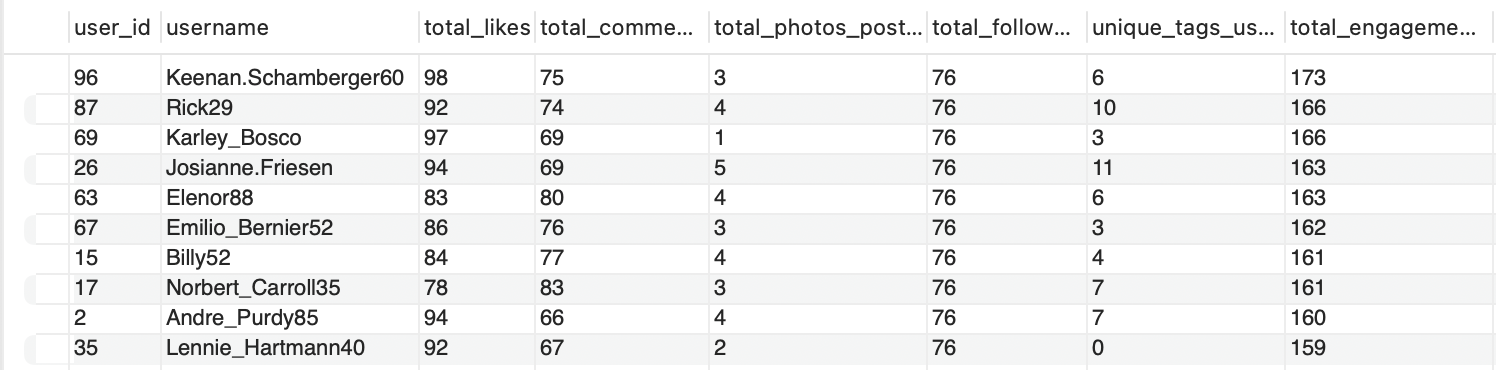


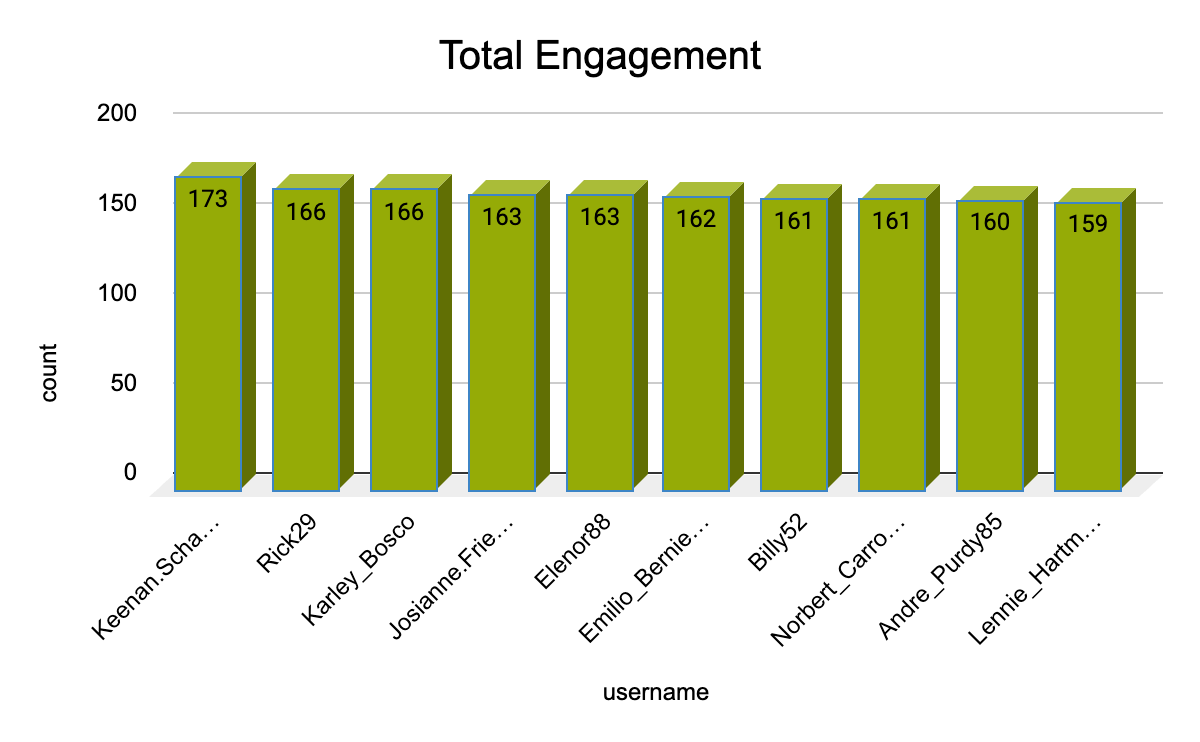
Output:

# **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**





To identify the most loyal and valuable users on a website, it is essential to analyze their engagement patterns. This can be achieved by tracking various user activities such as posting, liking, commenting, and participating in website-specific events.

**Query Development: A CTE-Based Approach:**

To accomplish this task, we developed a query that leverages Common Table Expressions (CTEs) and joins to extract relevant data from multiple sources. The query consists of several stages:

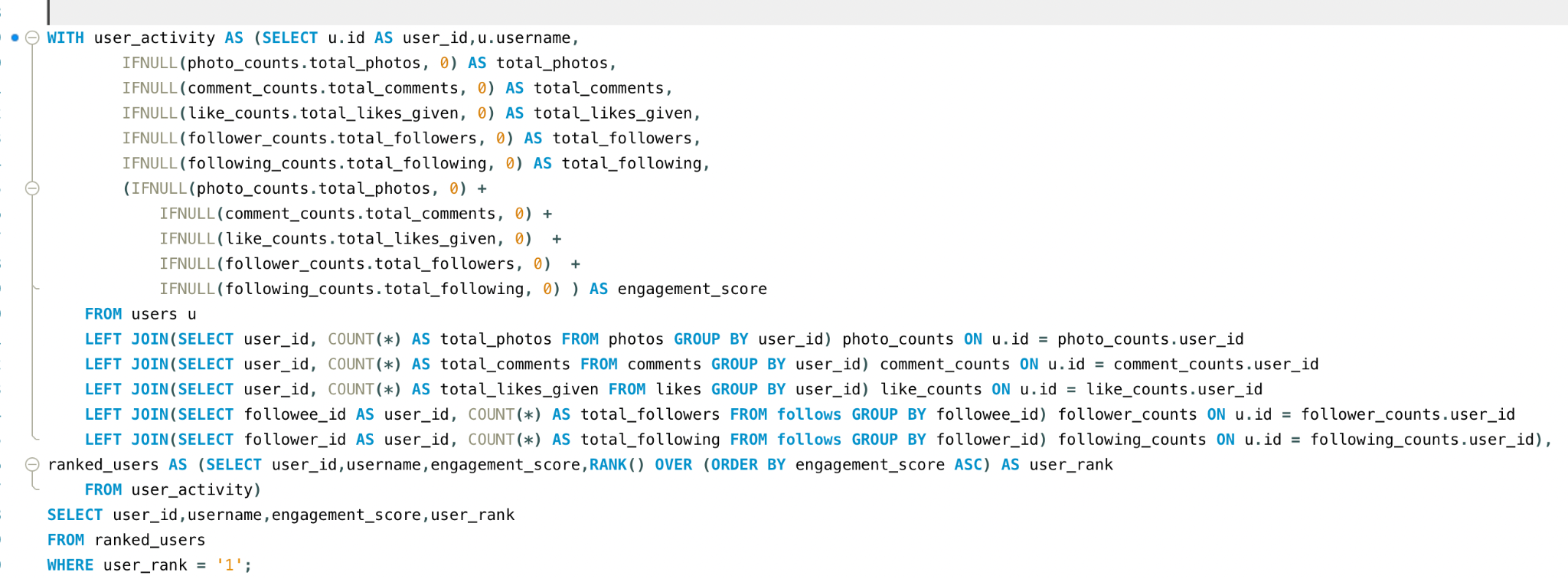
1. **User Post Extraction:** Identify users with their total post counts.
2. **User Like Extraction:** Identify users with their total like counts.
3. **User Comment Extraction:** Identify users with their total comment counts.
4. **Total Engagement Calculation:** Calculate the total engagement for each user by summing their like and comment counts.
5. **Ranking and Filtering:** Order the results in descending order based on total engagement and filter the top 10 users.

**Insights and Recommendations:**

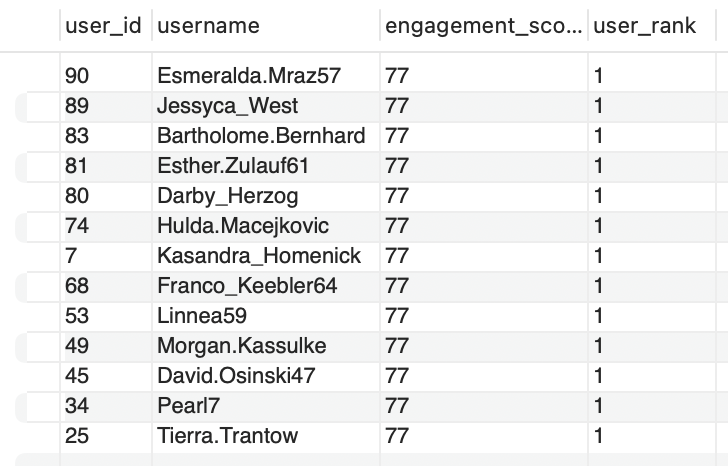
The resulting output provides valuable insights into the most loyal and valuable users on the website. These users can be further analyzed to understand their behavior, preferences, and interests. By recognizing and rewarding these users, the website can foster a sense of community, encourage user retention, and drive engagement.

**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 an engagement score for each user based on their activity (photos, comments, likes, followers, and following). It then ranks users from least to most engaged using the RANK() function. Finally, it selects the least engaged user (ranked 1) from the results.

**Output:**

To reinvigorate inactive users and encourage them to re-establish their engagement, we've developed a comprehensive strategy that incorporates personalized communication, gamification, content inspiration, and incentives.

**PERSONALIZED COMMUNICATION**

* **Dynamic Email Campaigns:** Design and deploy targeted email campaigns that leverage user-specific data to remind them of their previous activity, showcase what they've missed, and highlight new features or updates.
* **Real-Time Notifications:** Implement real-time notifications that alert users to relevant events, such as new posts from friends or mentions in comments.

**GAMIFICATION**

* **Challenges and Quests:** Introduce challenges and quests that incentivize users to engage with the platform, such as "Complete 5 posts in 7 days" or "Participate in 3 discussions this week."
* **Reward Systems:** Develop a reward system that acknowledges and rewards users for their achievements, such as badges, points, or leaderboard rankings.

**CONTENT INSPIRATION**

* **AI-Driven Recommendations:** Leverage AI-powered algorithms to provide users with personalized content recommendations based on their interests, preferences, and past behavior.
* **Trending Topics:** Highlight trending topics and popular discussions to inspire users to create new content or participate in existing conversations.

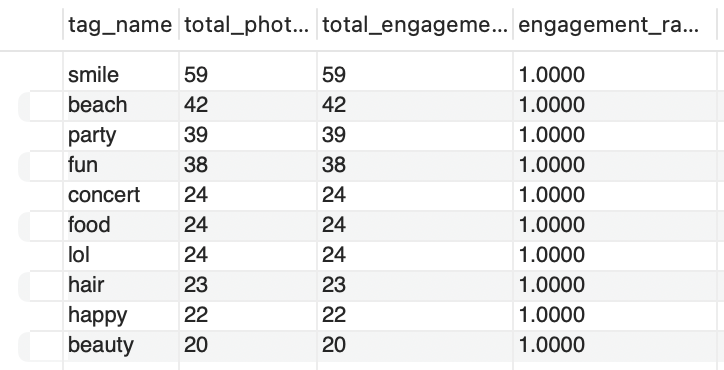
**INCENTIVES**

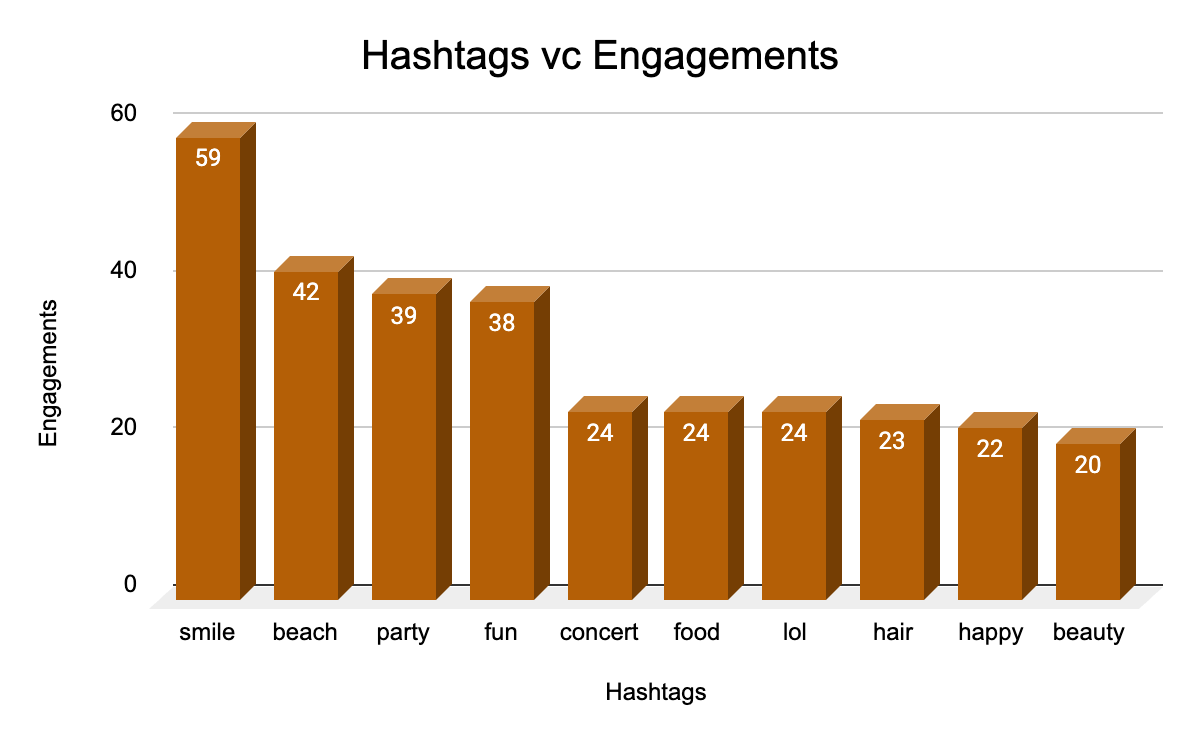
* **Exclusive Offers:** Provide exclusive offers, discounts, or promotions to users who return to the platform after a period of inactivity.
* **Premium Features:** Offer premium features or access to exclusive content for users who consistently engage with the platform over a set period.

**By implementing these strategies, we can effectively re-engage inactive users, encourage them to start posting or engaging again, and ultimately drive growth and retention on the platform.**

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

**Ans.**





To maximize user engagement and drive campaign success, it's essential to identify and leverage high-engagement hashtags. These hashtags are associated with posts that have garnered significant likes, comments, and shares, making them a valuable asset for content strategy and ad campaigns.

**IDENTIFYING HIGH-ENGAGEMENT HASHTAGS**

* **Data Analysis:** Conduct thorough data analysis to uncover hashtags linked to top-performing posts.
* **Hashtag Tracking:** Monitor hashtag performance over time to identify trends and patterns.

**CONTENT STRATEGY**

* **Challenges and Prompts:** Encourage users to incorporate high-engagement hashtags into their content by featuring them in challenges or prompts.
* **Campaign Development:** Create comprehensive campaigns centered around trending topics or popular hashtags.
* **Influencer Partnerships:** Collaborate with influencers who have a track record of using high-engagement hashtags to amplify campaign reach.

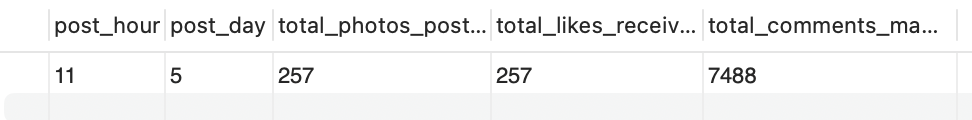
**AD CAMPAIGNS**

* **Targeted Advertising:** Utilize high-engagement hashtags to target ads, ensuring they align with popular topics and resonate with the interests and engagement patterns of your audience.
* **Ad Creative Development:** Design ad creative that incorporates high-engagement hashtags, increasing relevance and engagement.
* **Budget Allocation:** Allocate budget to ad campaigns that feature high-engagement hashtags, maximizing ROI and campaign effectiveness.

**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.**

For this I constructed a Query to check the total photos posted, total likes received on that photo, and total comments on those photos at particular time and days of week.



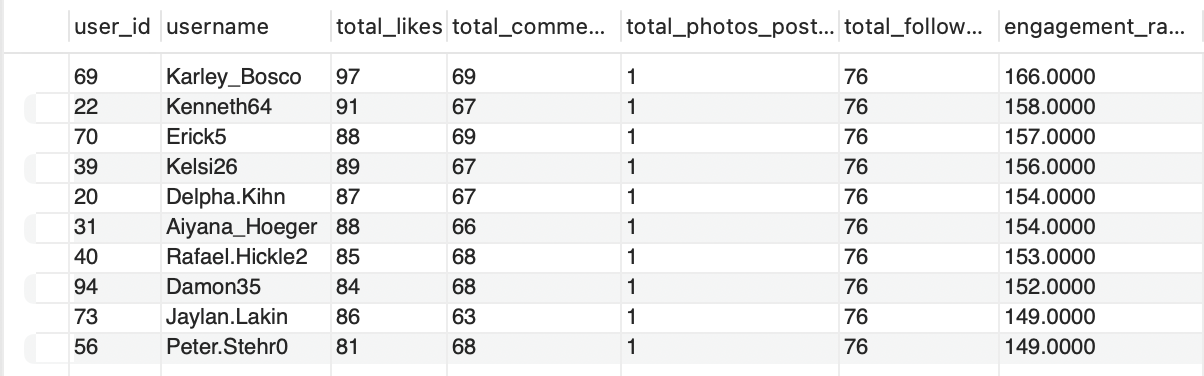
Through a meticulous analysis of user engagement patterns, we've discovered that Wednesday is the most opportune day to post, as it yields the highest response and engagement rates. This finding underscores the existence of a correlation between posting times and user engagement.

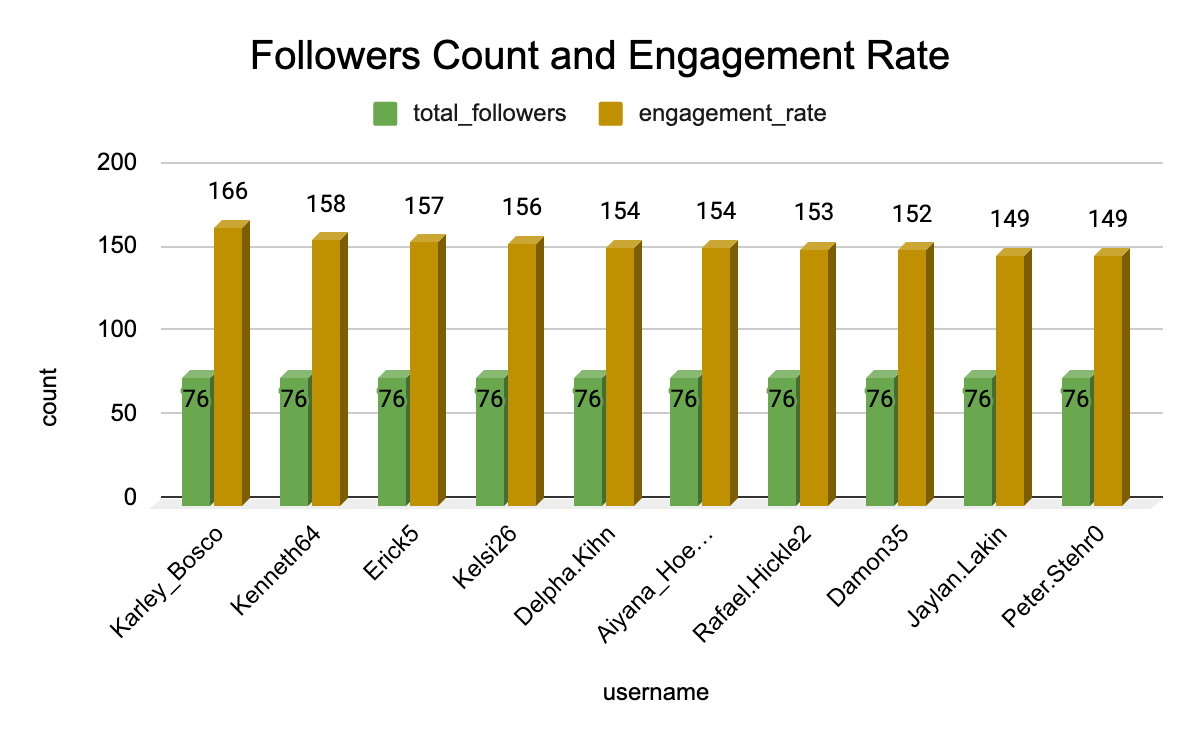
**Methodology**:

* **Temporal Analysis:** Conducted an exhaustive examination of time-based data to identify correlations between user demographics and engagement metrics. Pinpointed peak engagement periods across diverse demographic groups.
* **Targeted Campaigns:** Utilized insights to craft targeted advertisements and content. For instance, if younger users exhibit heightened activity during evening hours, scheduled posts and ads to appear during those times.
* **Content Personalization:** Created content that resonates with specific demographics based on their engagement patterns. Tailored messaging and imagery to align with the preferences of distinct groups.

**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.**





To pinpoint the ideal candidates for marketing campaigns, we first need to identify users with the highest engagement rates. To achieve this, I constructed a query that leverages Common Table Expressions (CTEs) to isolate the top 10 users with the most followers and subsequently, the highest engagement rates.

**Query Construction:**

* **CTE 1:** Top Followers: Identify users with the highest number of followers.
* **CTE 2:** Engagement Rates: Calculate total likes, comments, and engagement rates for each user in CTE 1.
* **Final Query:** Limit the results to the top 10 users with the highest engagement rates.

**Influential User Identification:**

By executing this query, we can accurately identify the top 10 influential users who have the highest engagement rates and are most likely to be effective brand ambassadors. These users can be leveraged to promote products, services, or ideas to their engaged audience, ultimately driving business growth and success.

To optimize user engagement and drive business growth, we employed a comprehensive strategy that involves segmenting users based on their engagement levels, content preferences, and time spent on the website

**SEGMENTATION CRITERIA**

* **Engagement-Based Segmentation**: Categorize users into distinct segments based on their engagement patterns, such as frequent likers, commenters, or content creators.
* **Content-Driven Segmentation**: Segment users based on their content preferences, including topics, formats, and interests.
* **Time-Based Segmentation**: Divide users into segments based on their time spent on the website, including frequency, duration, and recency of visits.

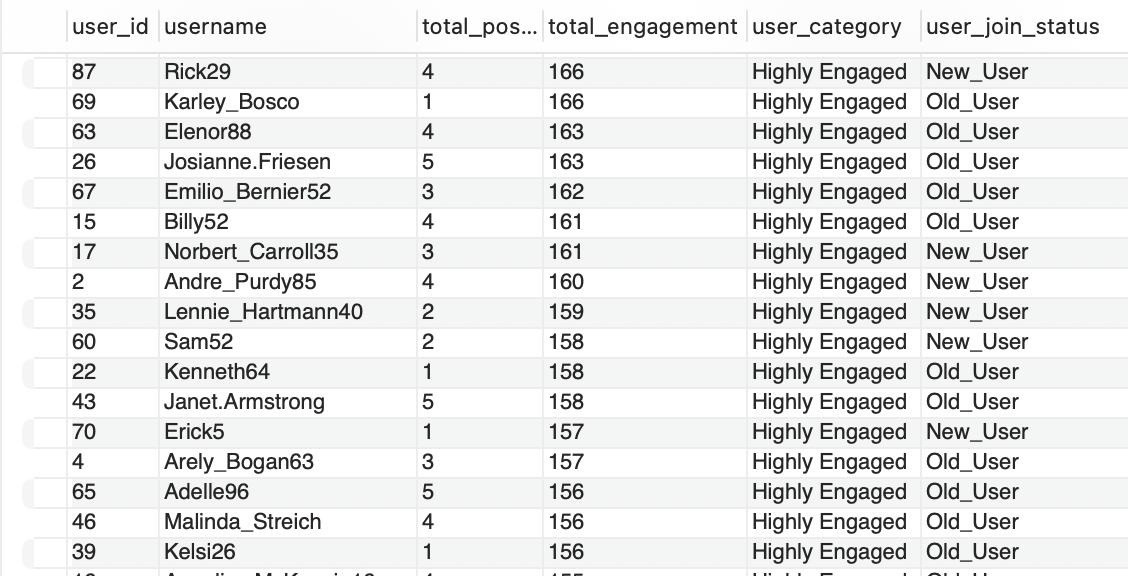
**TARGETED MARKETING**

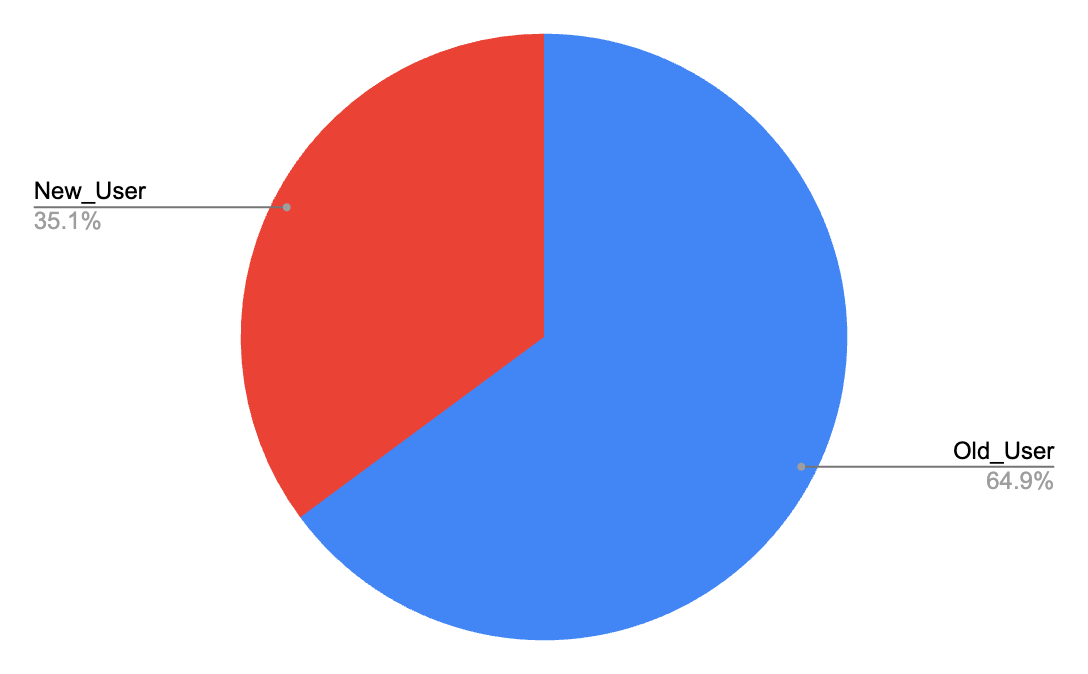
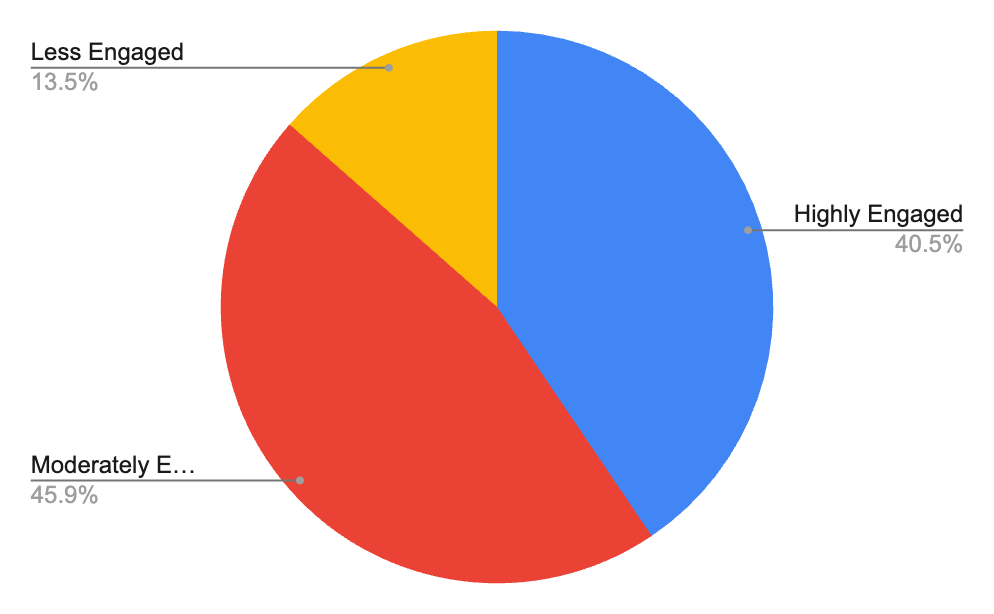
* **Segment-Specific Campaigns**: Develop targeted marketing campaigns tailored to each user segment, incorporating relevant messaging, imagery, and calls-to-action.
* **Personalized Recommendations**: Offer personalized recommendations for products, services, or content based on individual user preferences and behavior.
* **Dynamic Content**: Create dynamic content that adapts to user interests, engagement patterns, and demographics.

**PERSONALIZATION**

* **User-Centric Experience**: Design a user-centric experience that incorporates personalized elements, such as customized feeds, content suggestions, and targeted notifications.
* **Real-Time Engagement**: Foster real-time engagement through live updates, push notifications, and in-app messaging.

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

**Ans.** 



To develop effective marketing strategies, it's essential to understand user behavior and preferences. By segmenting users based on their engagement levels and account creation year, we can create targeted campaigns that resonate with specific groups.

**Segmentation Criteria:**

* **Engagement-Based Segmentation**: Classify users into three segments based on their total engagements:
  + Highly Engaged (total engagements > 150)
  + Moderately Engaged (total engagements between 100 and 150)
  + Less Engaged (total engagements < 100)
* **Account Creation Year-Based Segmentation**: Divide users into two segments based on their account creation year:
  + New Users (accounts created in or after 2017)
  + Old Users (accounts created before 2017)

**Query Construction:**

To achieve this segmentation, I employed a query that leverages Common Table Expressions (CTEs) and joins to extract relevant data. The query calculates total likes, comments, and engagements for each user and then uses the CASE function to classify them based on their engagement levels and account creation year.

**Results:**

The query yielded a list of users with their corresponding segments, providing valuable insights into user behavior and preferences. This information can be used to develop targeted marketing campaigns, personalize user experiences, and ultimately drive business growth.

To optimize user engagement and drive business growth, we employed a comprehensive strategy that involves segmenting users based on their behavior, preferences, and time spent on the website. This approach enables us to create targeted marketing campaigns, personalize user experiences, and ultimately enhance overall platform performance.

**SEGMENTATION CRITERIA**

* **Behavioral Segmentation**: Divide users into distinct segments based on their behavior, such as frequent likers, commenters, or content creators.
* **Preference-Based Segmentation**: Segment users based on their content preferences, including topics, formats, and interests.
* **Time-Based Segmentation**: Categorize users based on their time spent on the website, including frequency, duration, and recency of visits.

**TARGETED MARKETING**

* **Segment-Specific Campaigns**: Develop targeted marketing campaigns tailored to each user segment, incorporating relevant messaging, imagery, and calls-to-action.
* **Personalized Recommendations**: Offer personalized recommendations for products, services, or content based on individual user preferences and behavior.
* **Exclusive Content**: Provide exclusive content or offers to specific user segments, enhancing their engagement and loyalty.

**PERSONALIZATION**

* **Dynamic User Experiences**: Create dynamic user experiences that adapt to individual preferences, behavior, and interests.
* **Real-Time Engagement**: Foster real-time engagement through live updates, push notifications, and in-app messaging.
* **Continuous Improvement**: Continuously monitor user behavior and adjust personalization strategies to ensure optimal engagement and satisfaction.

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

**Ans.**

To assess the effectiveness of ad campaigns, it is essential to examine a range of metrics that provide insights into their performance in reaching and engaging the target audience. The objective is not only to evaluate the campaign's efficacy but also to utilize this information to optimize future campaigns for improved results.

**METRICS ANALYSIS**

* **Impressions**: High impressions indicate good visibility; however, they must be analyzed in conjunction with other metrics such as clicks and conversions to gain a comprehensive understanding of campaign performance.
* **Clicks**: A higher Click Rate suggests that the ad content, creative, and targeting are compelling and relevant to the audience. Conversely, a low Click Rate may indicate that the ad is not engaging or that it is being shown to the wrong audience.
* **Conversion Rate**: A high conversion rate indicates that users are not only interested enough to click the ad, but they are also motivated to take further action. Conversely, a low conversion rate could point to issues with the landing page, the offer, or the overall user experience.
* **Engagement Metrics**: High engagement indicates that the ad content resonates with the audience, which can increase brand awareness and loyalty, even if it doesn't immediately lead to conversions.

**OPTIMIZING FUTURE CAMPAIGNS**

* **Testing and Iteration**: Continuously test different versions of your ad (e.g., headlines, images, calls to action) to identify which performs better. Utilize the results to refine future ads and optimize campaign performance.
* **Targeting Refining**: Analyze which audience segments (age, gender, location, interests) are responding best to your ads. Adjust targeting to focus on high-performing segments, ensuring maximum ROI and campaign effectiveness.
* **Platform Optimization**: Evaluate which platforms or ad placements (e.g., social media, search engines, websites) are generating the best results. Allocate ad spend accordingly, maximizing visibility during peak times or locations.
* **Retargeting Strategies**: Implement retargeting to show ads to users who have previously interacted with your brand but didn’t convert. Utilize personalized ads to encourage them to complete the conversion, increasing overall campaign effectiveness.
* **Continuous Monitoring and Optimization**: Regularly monitor campaign performance and make adjustments as needed. Analyze data to identify areas for improvement, and implement changes to optimize campaign performance and maximize ROI.

**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**.

To effectively identify potential brand ambassadors or advocates from a pool of users with varying engagement levels—Highly Engaged, Moderately Engaged, and Less Engaged—we can use a strategic approach that evaluates a combination of factors related to user activity, reach, content quality, and authenticity. Here’s how we can leverage user activity data to pinpoint those individuals who could represent the brand in a meaningful and impactful way:

### **1. Identifying Highly Engaged Users:**

The first step is to track users who exhibit consistently high engagement across multiple forms of interaction—such as likes, comments, shares, and responses to stories or posts. These individuals not only interact frequently with content but also demonstrate a deeper level of interest and connection to the brand. Their consistent engagement indicates that they are genuinely involved and invested in the brand’s presence.

Beyond simple engagement, it’s crucial to focus on users who actively share the brand’s content or mention the brand in their posts and stories. Sharing brand content amplifies the brand’s reach and signals that the user is not just interacting but also promoting the brand’s message to a broader audience, making them more likely to serve as organic advocates.

### **2. Analyzing Influence and Reach:**

To assess the potential impact of users, it’s important to look at their follower count and overall reach. While follower count alone is not a definitive indicator, it serves as an initial filter to identify users who have the potential to influence a wider audience. A larger follower base increases the likelihood that the user’s posts will be seen by more people, and if they align with the brand’s target audience, this can create significant opportunities for brand exposure.

Additionally, it is essential to assess whether the user is already an influencer within a specific niche. For example, if the brand is focused on a particular industry, such as fitness, wellness, or fashion, identifying users who are already seen as thought leaders or influencers within that niche can provide valuable insight. Influencers who already embody the values and interests of the brand can be powerful advocates, as their recommendations will likely carry more weight within their communities.

### **3. Evaluating Content Quality and Alignment with Brand Values:**

The next step is to examine the type of content the user regularly posts to determine whether it aligns with the brand’s image and values. Content creators who consistently produce material that resonates with the brand’s core messaging are more likely to serve as effective ambassadors. For instance, if the brand is promoting an eco-friendly product, users who share content around sustainability, environmental issues, or green living could be a perfect fit for representing the brand’s initiatives.

Additionally, content quality plays a crucial role. Users who create high-quality, visually appealing, and original content can help elevate the brand’s image when they share it. The aesthetics, creativity, and production value of their posts can reflect positively on the brand, enhancing its overall reputation.

### **4. Tracking Loyalty and Long-Term Engagement:**

Beyond short-term interactions, it is important to identify users who have demonstrated consistent and long-term engagement with the brand. Long-term loyalty indicates a deeper, more genuine connection with the brand, making these users ideal candidates for ambassadorship roles. Look for users who have been engaging with the brand for an extended period, not just during promotional campaigns or giveaways.

Another critical factor is identifying users who have voluntarily advocated for the brand in the past. Users who have recommended or praised the brand without incentives—whether through organic testimonials, reviews, or social media posts—show that their support is genuine. Such users are more likely to continue advocating for the brand in the future, even in a formal ambassador role.

### **5. Evaluating Authenticity and Advocacy Potential:**

Authenticity is a cornerstone of effective brand advocacy. It’s essential to assess whether a user’s engagement with the brand appears to be driven by genuine interest or if they are simply participating for incentives (e.g., giveaways or free products). True brand advocates naturally align with the brand’s values and communicate their experiences in a way that resonates with their audience.

A powerful indicator of authenticity is user-generated content (UGC) where individuals share personal experiences with the brand’s products or services. These types of testimonials—whether in the form of reviews, unboxing videos, or social media posts—serve as proof of genuine engagement and can be a compelling reason to bring these users on board as brand ambassadors. Their real, unfiltered stories make their advocacy more credible and relatable to potential customers.

### **6. Engaging with Potential Ambassadors:**

Once these high-potential users have been identified, it’s essential to engage with them directly to build a relationship and gauge their willingness to collaborate. This engagement can take the form of personalized comments, likes, or even direct messages, showing them that the brand recognizes their contributions and values their involvement.

After establishing rapport, brands can offer these users exclusive opportunities to participate in special events, beta tests, product launches, or other brand initiatives. Providing them with unique access not only strengthens the relationship but also gives them a platform to organically promote the brand. These interactions can help foster a long-term partnership, turning engaged users into loyal ambassadors who actively advocate for the brand.

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

**Ans**.

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 analyze and come up with insights from the given data.

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

### **1. Problem Identification:**

My first step would be to fully grasp the objectives of the analysis—whether the goal is to optimize user engagement, boost retention, identify potential brand ambassadors, or something else. Understanding the specific goals will allow me to devise a clear plan and approach for how to manipulate and analyze the data effectively to achieve the desired outcome.

### **2. Data Exploration:**

Once the problem is defined, I would begin with an exploratory analysis of the dataset. This involves reviewing the data structure, distributions, and identifying any initial patterns or anomalies. At this stage, I would also determine whether any meaningful insights can be drawn directly from the raw data or if further data processing is required to uncover trends.

### **3. Data Cleaning and Validation:**

Next, I would clean the data. This step includes checking for null values, duplicates, or any inconsistencies across the tables. If any such issues are found, I would correct them to ensure that the data is reliable and ready for deeper analysis. Data validation ensures that the analysis is based on accurate and clean information.

### **4. Trend Formation:**

With clean data in hand, I would start forming hypotheses around possible trends. For example, one might hypothesize that users who engage more frequently with content are more likely to retain over time. I would then analyze related metrics—such as user activity, follower growth, content interactions, and more—to test these hypotheses and validate whether there is a correlation with the desired business goals.

### **5. Data-Driven Solutions:**

At this stage, I would perform SQL queries to delve deeper into the data and verify if the initial hypotheses hold true. This is where I would test my assumptions and analyze the actual trends in the data. If the initial hypotheses don't align with the data, I would identify where the discrepancies lie and investigate how the current trends differ from what was expected, refining the approach if needed.

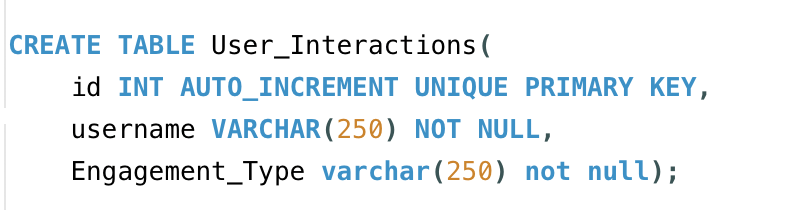
### **6. Insights from the Current Scenario:**

Finally, I would compile the results from the SQL queries into a comprehensive report. This report would provide a clear picture of the current state of affairs, highlighting areas where performance is strong and identifying where improvements are needed. Based on these insights, I would suggest actionable recommendations to address the gaps, providing management with the necessary data to make informed decisions and take corrective actions where required.

**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?

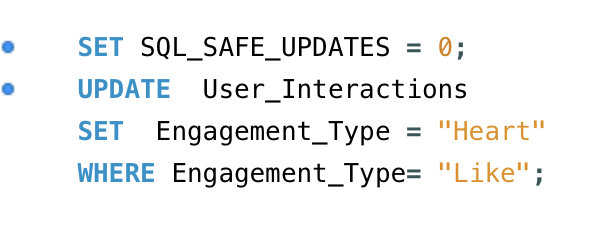
**Ans.**

For this problem first I created a new Table User\_Interactions with the following Query.



After creating the User\_Interaction table then I entered values in the table taking from the original table data.



Finally for updating the values in the table User\_Interaction which was the original problem statement I constructed the following SQL Query  
  


From the above query I got the desired output updating the "Engagement\_Type" column to change all instances of "Like" to "Heart" to align with Instagram's terminology.

