**Objective Questions**

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

**for checking if the dataset contains duplicate values or not, we can use these queries on all given tables:**

* By running below queries we got to know there are no duplicates present in tables.

**For checking duplicates present in tables:**

select id as user\_id ,

count(\*) as duplicate\_values\_in\_users

from users

group by id

having duplicate\_values\_in\_users>1;

select id as tag\_id,

count(\*) as duplicate\_values\_in\_tags

from tags

group by id

having duplicate\_values\_in\_tags>1;

select id as photo\_id,

count(\*) as duplicate\_values\_in\_photos

from photos

group by id

having duplicate\_values\_in\_photos>1;

select photo\_id,

tag\_id,

count(\*) as duplicate\_values\_in\_phototags

from photo\_tags

group by photo\_id,tag\_id

having duplicate\_values\_in\_phototags >1;

select user\_id,

photo\_id,

count(\*) as duplicate\_values\_in\_likes

from likes

group by user\_id,photo\_id

having duplicate\_values\_in\_likes>1;

select id as comment\_id ,

count(\*) as duplicate\_values\_in\_comments

from comments

group by id

having duplicate\_values\_in\_comments>1;

select follower\_id

,followee\_id,

count(\*) as duplicate\_values\_in\_follows

from follows

group by follower\_id,followee\_id

having duplicate\_values\_in\_follows >1;

* By running below queries we got to know there are no null values present in tables.

**For checking null values present in tables:**

select count(\*) as null\_values\_in\_users

from users

where id is null or username is null or created\_at is null;

select count(\*) as null\_values\_in\_tags

from tags

where id is null or tag\_name is null or created\_at is null;

select count(\*) as null\_values\_in\_photos

from photos

where id is null or image\_url is null or user\_id is null or created\_dat is null;

select count(\*) as null\_values\_in\_photo\_tags

from photo\_tags

where photo\_id is null or tag\_id is null;

select count(\*) as null\_values\_in\_follows

from follows

where follower\_id is null or followee\_id is null or created\_at is null;

select count(\*) as null\_values\_in\_comments

from comments

where id is null or comment\_text is null or user\_id is null or photo\_id is null or created\_at is null;

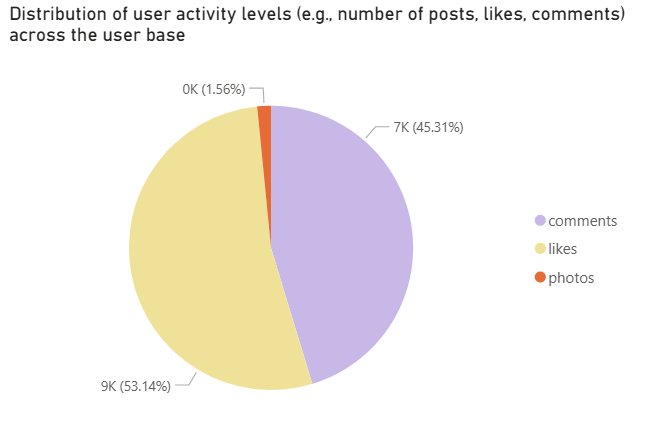
select count(\*) as null\_values\_in\_likes

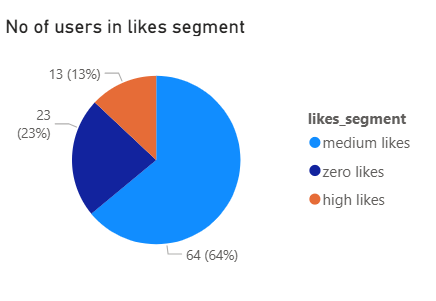
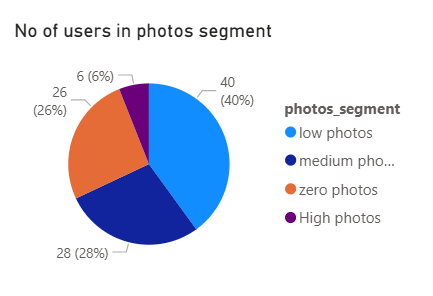
from likes

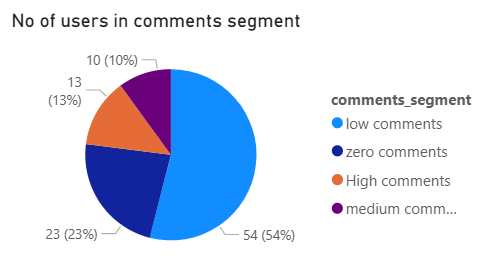
where user\_id is null or photo\_id is null or created\_at is null;

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

The query calculates and categorizes user activity levels across the platform based on their number of likes, comments, and photo posts. Each user is segmented into one of four activity categories — zero, low, medium, or high — for each metric. This segmentation is done by comparing individual user counts with the overall distribution (min and max values) for each activity type. The output provides a comprehensive view of each user's engagement level, helping identify highly active users, moderately engaged users, and inactive or minimally active users across the platform.





The below query was used to retrieve the information and then the result was sent to powerbi In powerbi, result was used to create these charts ; the above doughnut charts.

with count\_likes as(

select user\_id,

count(\*) as num\_likes

from likes

group by user\_id

)

, count\_comments as(

select user\_id,

count(\*) as num\_comments

from comments

group by user\_id

)

, count\_photos as(

select user\_id,

count(\*) as num\_photos

from photos

group by user\_id

)

select a.id,

a.username,

coalesce(num\_likes,0) as num\_likes,

case

when coalesce(num\_likes,0) between 1 and ((select max(num\_likes) from count\_likes)-(select min(num\_likes) from count\_likes))/3 then "less likes"

when coalesce(num\_likes,0)>2\*((select max(num\_likes) from count\_likes)-(select min(num\_likes) from count\_likes))/3 then "high likes"

when coalesce(num\_likes,0) =0 then "zero likes"

else "medium likes" end as likes\_segment

,coalesce(num\_comments,0) as num\_comments,

case

when coalesce(num\_comments,0) between 1 and ((select max(num\_comments) from count\_comments)-(select min(num\_comments) from count\_comments))/3 then "low comments"

when coalesce(num\_comments,0)>2\*((select max(num\_comments) from count\_comments)-(select min(num\_comments) from count\_comments))/3 then "High comments"

when coalesce(num\_comments,0)=0 then "zero comments"

else "medium comments" end as comments\_segment,

coalesce(num\_photos,0) as num\_photos ,

case

when coalesce(num\_photos,0) between 1 and ((select max(num\_photos) from count\_photos)-(select min(num\_photos) from count\_photos))/3 then "low photos"

when coalesce(num\_photos,0)>2\*((select max(num\_photos) from count\_photos)-(select min(num\_photos) from count\_photos))/3 then "High photos"

when coalesce(num\_photos,0)=0 then "zero photos"

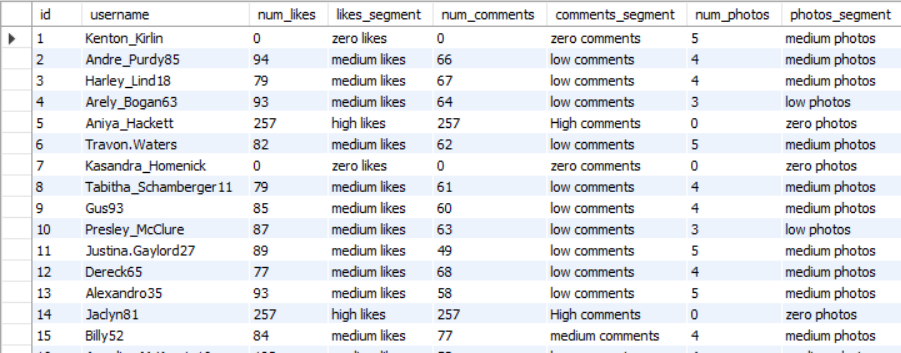
else "medium photos" end as photos\_segment

from users a

left join count\_likes b on a.id=b.user\_id

left join count\_comments c on a.id=c.user\_id

left join count\_photos d on a.id=d.user\_id;



**Insights** :

* Majority of users fall into the low or medium engagement categories across likes, comments, and photo uploads.
* High-engagement users represent a smaller fraction of the user base. This suggests a lack of consistent high-contribution behavior across the platform.
* A significant portion of users fall in the zero activity category, indicating either limited awareness, usability issues, or lack of motivation to engage.

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



select round(avg(num\_tags),2) as avg\_num\_tags\_per\_post

from (

select a.id,

coalesce(count(tag\_id),0) as num\_tags

from photos a

left join photo\_tags b on a.id=b.photo\_id

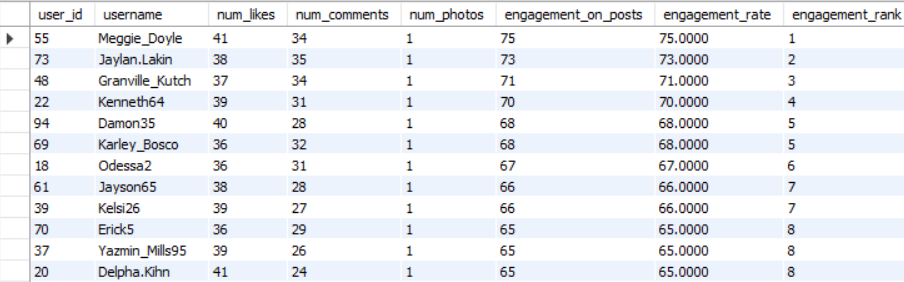
group by a.id

) as ab;

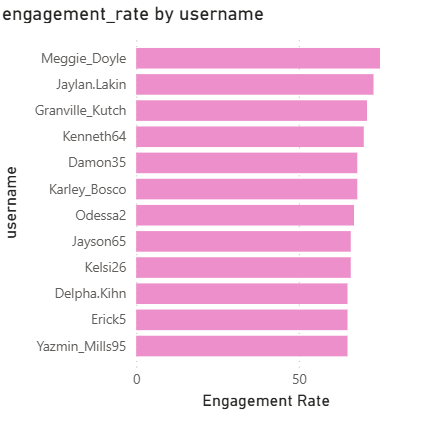
**4)Identify the top users with the highest engagement rates (likes, comments) on their posts and rank them.**

We defined engagement rate as no of likes and comments that a user has across all its posts :

Engagement rate= (number of likes + number of comments)/number of photos



The table above shows the engagement rate (likes + comments)/photos and we have ranked users based on the engagement rate in the ‘engagement\_rank’ column. We have taken the highest engagement rate (ranking up to 8 ).



The query we used to get output is:

with post\_engagement as

(select a.user\_id ,

d.username ,

count(distinct b.user\_id) as num\_likes,

count(distinct c.id) as num\_comments,

count(distinct a.id) as num\_photos

,count(distinct b.user\_id)+count(distinct c.id) as engagement\_on\_posts

from photos a

left join likes b on a.id=b.photo\_id

left join comments c on a.id=c.photo\_id

join users d on a.user\_id=d.id

group by user\_id)

,engagement\_rate as

(select \*,

case

when num\_photos=0 then 0 else

(num\_likes+num\_comments)/num\_photos

end as engagement\_rate

from post\_engagement)

select \*,

dense\_rank()over(order by engagement\_rate desc) as engagement\_rank

from engagement\_rate;

**Insights:**

Meggie\_doyle was founded to be the user with highest engagement rate of 75. These users are to be considered in frequent users category, we can tag them as “Frequent User” and give them features like:

* Beta feature rollouts.
* Early access to new tools or filters.
* Invitations to creator/influencer programs.

**5)Which users have the highest number of followers and followings?**

The query we used to get the output is:

with followings as(

select follower\_id as user\_id

,count(\*) as number\_of\_followings

from follows

group by follower\_id

)

,followers as(

select followee\_id as user\_id

,count(\*) as number\_of\_followers

from follows

group by followee\_id

)

select a.user\_id

,c.username

,a.number\_of\_followings

,b.number\_of\_followers

from followings a

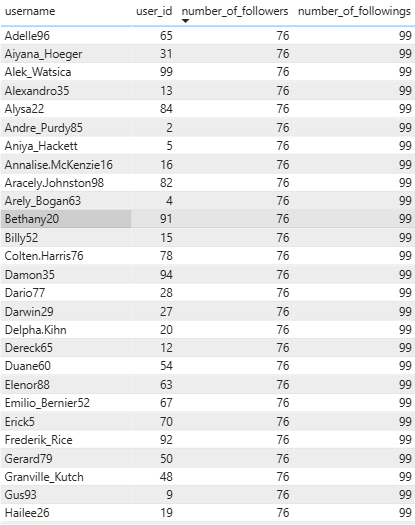
join followers b on a.user\_id=b.user\_id

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

where number\_of\_followings=(select max(number\_of\_followings) from followings)

or number\_of\_followers=(select max(number\_of\_followers) from followers)

The below table shows name of users with highest followers and highest following altogether are:



**Insights**:

* They are likely very active socially, building networks, and possibly engaging in reciprocal behavior (follow-backs).
* Since they have both high followers and followings, these users are ideal virality agents; perfect for boosting campaign reach.

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

This query calculates the average engagement rate per post for each user on the platform, where engagement rate is defined as:

Engagement Rate = (Number of Likes + Comments) / Number of Posts

The query we used to get the output is:

with post\_engagement as

(select a.user\_id ,

d.username ,

count(distinct b.user\_id) as num\_likes,

count(distinct c.id) as num\_comments,

count(distinct a.id) as num\_photos

,count(distinct b.user\_id)+count(distinct c.id) as engagement\_on\_posts

from photos a

left join likes b on a.id=b.photo\_id

left join comments c on a.id=c.photo\_id

join users d on a.user\_id=d.id

group by user\_id)

,engagement\_rate\_user as

(select \*,

case

when num\_photos=0 then 0

Else (num\_likes+num\_comments)/num\_photos end as engagement\_rate

from post\_engagement)

select user\_id,

username,

engagement\_on\_posts,

round(avg(engagement\_rate),2) as avg\_engagement\_rate

from engagement\_rate\_user

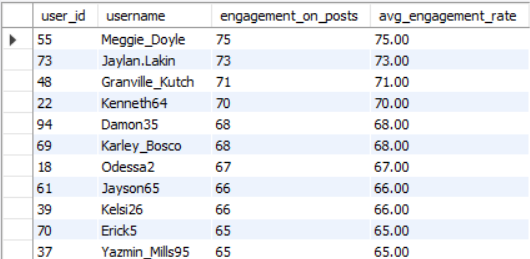
group by

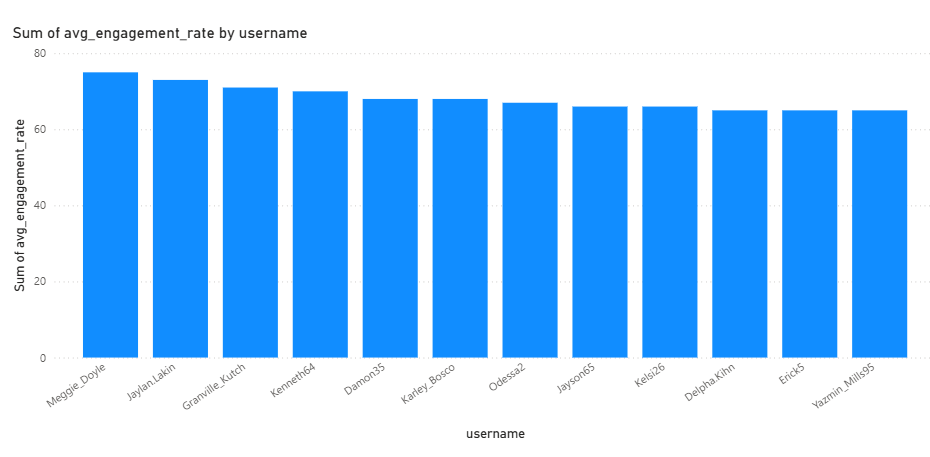
user\_id,

username

order by avg\_engagement\_rate desc;

the below table shows average engagement rate per post per users for top 10 users and bottom 10 users.





**Insights**:

* Users with the highest average engagement rates are producing quality content that resonates with others.
* A user with fewer posts can still rank high if their content gets a lot of interaction.
* Consistent high engagement suggests their posts may be driving trends, conversations, or community involvement.

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

The query we used to get the output is:

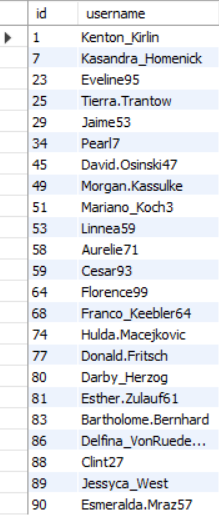
select a.id,

a.username from users a

left join likes b on a.id=b.user\_id

where b.user\_id is null;

The below table shows usernames i.e individuals who didn’t liked any posts:



**Insights**:

* Ther are 23 users who have never liked any posts.
* These users are active on the platform (i.e., they’ve created accounts), but have not interacted with content through likes.
* This behavior can imply:
* Passive usage – users scroll without engaging
* Lack of content relevance or interest
* New user behavior – just starting to explore
* Observer-type usage – platform used primarily for viewing

**8) How can you leverage user-generated content (posts, hashtags, photo tags) to**

**create more personalized and engaging ad campaigns?**

The query used to get the output –

with tag\_details as(

select d.id as user\_id,

tag\_id,

tag\_name,

count(distinct c.id) as num\_tags

from tags a

join photo\_tags b on a.id=b.tag\_id

join photos c on c.id=b.photo\_id

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

group by tag\_id,tag\_name,d.id

order by user\_id),

ranking\_tags as(

select \*,

dense\_rank()over(partition by user\_id order by num\_tags desc) as ranking

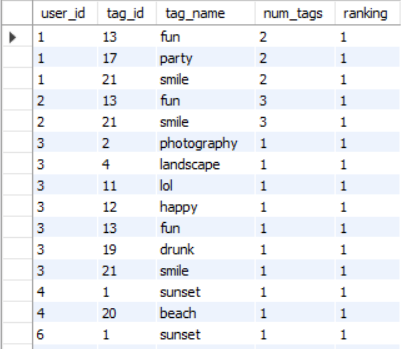
from tag\_details

)

select

\* from ranking\_tags where ranking=1;

The below table shows the user id, tag name and how many times that tag is used by a user. This will help us know what kind of tags user generally use and would give us an idea of his preferences.



**Insights**:

* Tag ‘Smile” is shown to be most used by all users across various posts and rule the trends game.

To Leverage This Data for Marketing & Ads

1. Tag-Based Segmentation for Personalized Ads

* Group users by their most-used tags.
* Deliver highly personalized campaigns.

2. Lookalike Audiences Based on Hashtag Use

* Identify top-tag users and create lookalike segments on platforms like Meta/Google Ads.

Run campaigns like:

* “Most Loved #PhotgraphOfTheDay” picks of the month
* “Tag your “smile” setup to win!”

When grouped and identifying the tags across users we can send personalized content to them regarding their tags(for this we used an extended query for grouping tags, which extends after the above query as:

with tag\_details as(

select d.id as user\_id,tag\_id,tag\_name,count(distinct c.id) as num\_tags from tags a join photo\_tags b on a.id=b.tag\_id join photos c on c.id=b.photo\_id

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

group by tag\_id,tag\_name,d.id

order by user\_id),

ranking\_tags as(

select \*,dense\_rank()over(partition by user\_id order by num\_tags desc) as ranking from tag\_details

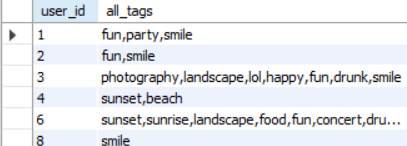
),

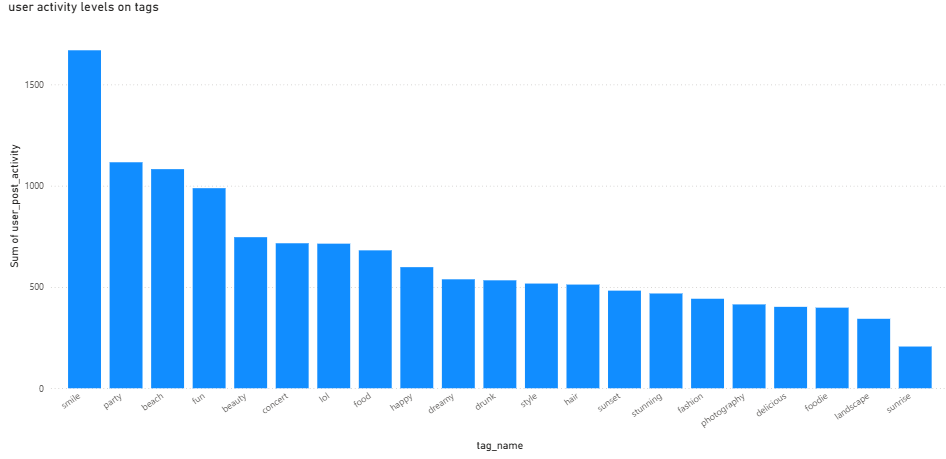
grouping\_tags as(

select

\* from ranking\_tags where ranking=1)

select user\_id,group\_concat(tag\_name) as all\_tags from grouping\_tags group by user\_id)





**9)Are there any correlations between user activity levels and specific content**

We have data on photos only. So,we can check the correlation between user activities on photos,

Using our query below:

with post\_data as(

select a.id as post\_id,

d.username,

count(distinct b.user\_id) +count(distinct c.id) as user\_post\_activity

from photos a

join likes b on a.id=b.photo\_id

join comments c on a.id=c.photo\_id

join users d on

a.user\_id=d.id

group by a.id)

,photo\_tag as(

select photo\_id,

tag\_name

from photo\_tags a

join tags b on b.id=a.tag\_id

)

select a.post\_id,

username,

tag\_name,

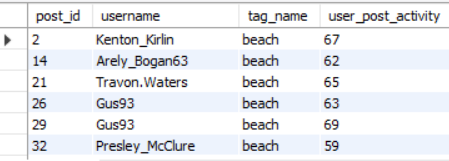
user\_post\_activity

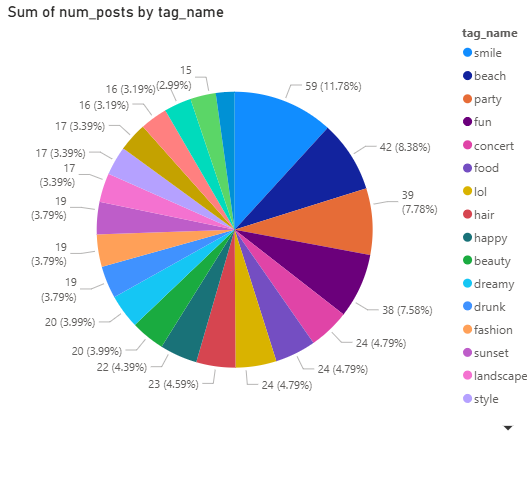
from post\_data a

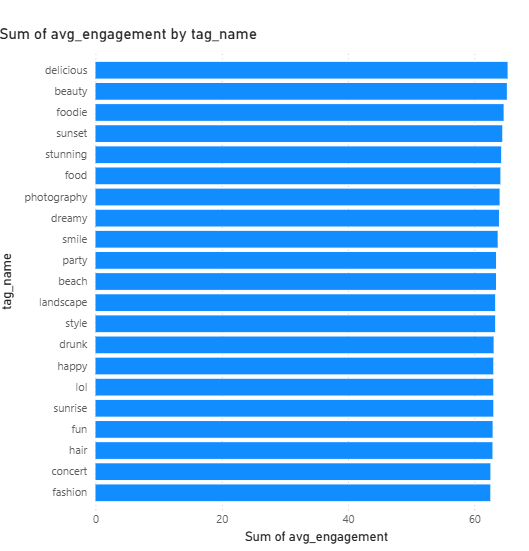
join photo\_tag b on a.post\_id=b.photo\_id

The below table explains how a post uploaded by a user with a certain tag is received by other users

Where user\_post\_activity=number of likes+number of comments:







**Insights**:

While smile is the most used tag out of all, ”delicious” tag has highest avg\_enagement of all tags(procured from extended query:

with post\_data as(

select a.id as post\_id,d.username,count(distinct b.user\_id) +

count(distinct c.id) as user\_post\_activity from photos a join likes b on a.id=b.photo\_id join comments c on a.id=c.photo\_id join users d on

a.user\_id=d.id

group by a.id)

,photo\_tag as(

select photo\_id,tag\_name from photo\_tags a join tags b on b.id=a.tag\_id

)

select

tag\_name,

avg(user\_post\_activity) as avg\_engagement,

count(\*) as num\_posts

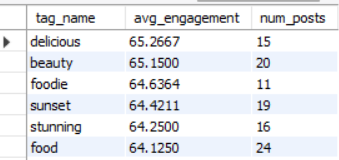
from (

select a.post\_id,username,tag\_name,user\_post\_activity from post\_data a join photo\_tag b on a.post\_id=b.photo\_id

) sub

group by tag\_name

order by avg\_engagement desc;



**Insights**:

* + Tag Popularity
  + The “smile” tag is the most frequently used across posts.
  + “Party” follows as the second most common tag.
  + Engagement Performance
  + The “delicious” tag, although used less frequently, generates the highest average engagement (likes + comments).
  + This suggests users respond more actively to food-related content.

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

with user\_post\_attributes as(

select a.id as user\_id,

a.username,

b.id as post\_id,

count(distinct c.user\_id) as no\_of\_likes ,

count(distinct d.id) as no\_of\_comments ,

count(distinct e.tag\_id) as no\_of\_tags

from users a

join photos b on a.id=b.user\_id

left join likes c on b.id=c.photo\_id

left join comments d on

b.id=d.photo\_id

join photo\_tags e on b.id=e.photo\_id

group by a.id,a.username,b.id

)

select user\_id,

username,

sum(no\_of\_likes) as total\_likes,

sum(no\_of\_comments) as total\_comments,

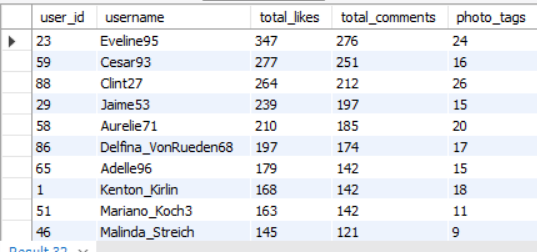
sum(no\_of\_tags) as photo\_tags from

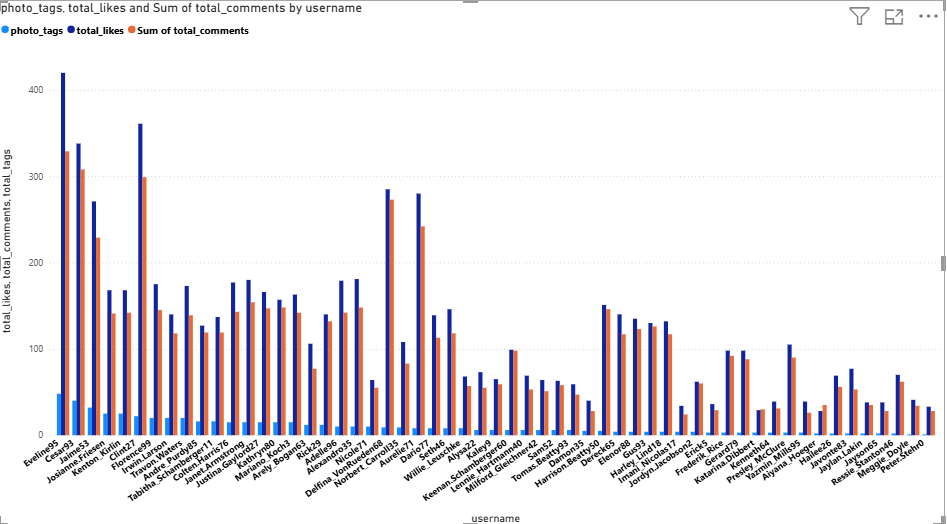
user\_post\_attributes

group by user\_id,

username

order by total\_likes+total\_comments+photo\_tags desc;





**Insights**:

Top Engaging Users by Likes

1. Eveline35 – Highest total likes across her posts
2. Jennings33 – Strong engagement through likes
3. Kenton\_Kirlin – Consistent like performance

High Comment-Generating Users

* Eveline35 and Joana\_Kemmer have significant community interaction via comments
* Indicates strong follower interest and active discussions

Balanced Performers

* Gus93 and Enrico\_Feil89 maintain a healthy like-to-comment ratio, showcasing well-rounded engagement
* Ideal user profiles for brand collaborations or beta content testing

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

with likes\_count as (

select

b.user\_id,

date\_format(b.created\_dat, '%Y-%m') as month,

count(distinct c.user\_id) as like\_count

from photos b

join likes c on b.id = c.photo\_id

group by b.user\_id, month

),

comments\_count as (

select

b.user\_id,

date\_format(b.created\_dat, '%Y-%m') as month,

count(distinct d.user\_id) as comment\_count

from photos b

join comments d on b.id = d.photo\_id

group by b.user\_id, month

),

engagement as (

select

a.id as user\_id,

a.username,

coalesce(lc.month, cc.month) as month,

coalesce(like\_count, 0) as like\_count,

coalesce(comment\_count, 0) as comment\_count,

coalesce(like\_count, 0) + coalesce(comment\_count, 0) as total\_engagement

from users a

left join likes\_count lc on a.id = lc.user\_id

left join comments\_count cc on a.id = cc.user\_id and lc.month = cc.month

)

select

user\_id,

username,

month,

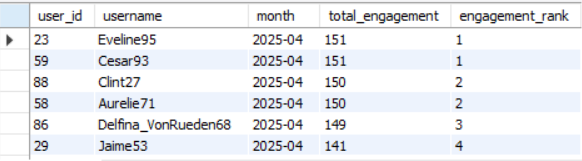
total\_engagement,

dense\_rank() over (partition by month order by total\_engagement desc) as engagement\_rank

from engagement

where month = '2025-04'

order by total\_engagement desc;



**Insights**:

* + Eveline95 and Cesar93 tied for Rank 1 with the highest engagement (151).
  + Clint27 and Aurelie71 are very close behind.
  + The top 6 users have very high engagement, indicating they could be influencers or highly active users.

**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.

This query helps grouping the tags in most liked pictures and sorting them in descending order:

with photo\_num\_likes as(

select photo\_id,

avg(total\_likes) as avg\_likes

from(select photo\_id,

count(user\_id) as total\_likes

from likes

group by photo\_id

)ab

group by photo\_id),

pic\_with\_highest\_avg\_likes as(

select photo\_id ,pic\_avg\_like\_rank

from(select \*,

dense\_rank()over(order by avg\_likes desc) as pic\_avg\_like\_rank

from photo\_num\_likes

)ab)

select a.photo\_id,group\_concat(tag\_name)as tags\_of\_the\_post

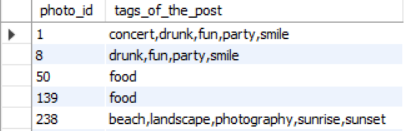
from photo\_tags a

join pic\_with\_highest\_avg\_likes b on a.photo\_id=b.photo\_id

join tags c on a.tag\_id=c.id

group by a.photo\_id

order by pic\_avg\_like\_rank desc;



**Insights**:

* These three tags (fun, party, smile) repeat in photo\_id 1 and 8 — which appear highest in rank.

Suggests that emotionally loaded, socially vibrant themes resonate best.

* The lesser visibility of “food” Suggests that “food-only” content underperforms, or lacks emotional/community pull compared to socially vibrant posts.

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

select

fb.followee\_id as original\_follower,

fb.follower\_id as user\_followed\_back

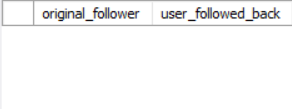
from follows fb

join follows orig

on fb.follower\_id = orig.followee\_id

and fb.followee\_id = orig.follower\_id

and fb.created\_at > orig.created\_at;



**Insights**:

The result of our query shows an empty table as the created\_at is same for all meaning everyone followed one another at same time.

**Subjective Questions**

**1)Based on user engagement and activity levels, which users would you consider the**

**most loyal or valuable? How would you reward or incentivize these users?**

with user\_likes as (

select user\_id,

count(photo\_id) as num\_likes

from likes

group by user\_id

),

user\_comments as (

select user\_id,

count(id) as num\_comments

from comments

group by user\_id

),

user\_photos as (

select user\_id,

count(id) as num\_photos

from photos

group by user\_id

)

,user\_table as(

select

a.id,

a.username,

coalesce(b.num\_likes, 0) + coalesce(c.num\_comments, 0) + coalesce(d.num\_photos, 0) as user\_activity,

coalesce(b.num\_likes, 0) + coalesce(c.num\_comments, 0) as user\_engagement,

dense\_rank() over (

order by

coalesce(b.num\_likes, 0) + coalesce(c.num\_comments, 0) + coalesce(d.num\_photos, 0) desc,

coalesce(b.num\_likes, 0) + coalesce(c.num\_comments, 0) desc

) as loyal\_customers\_ranking

from users a

left join user\_likes b on a.id = b.user\_id

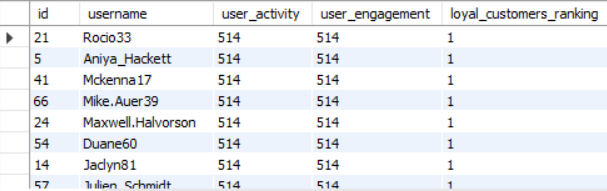
left join user\_comments c on a.id = c.user\_id

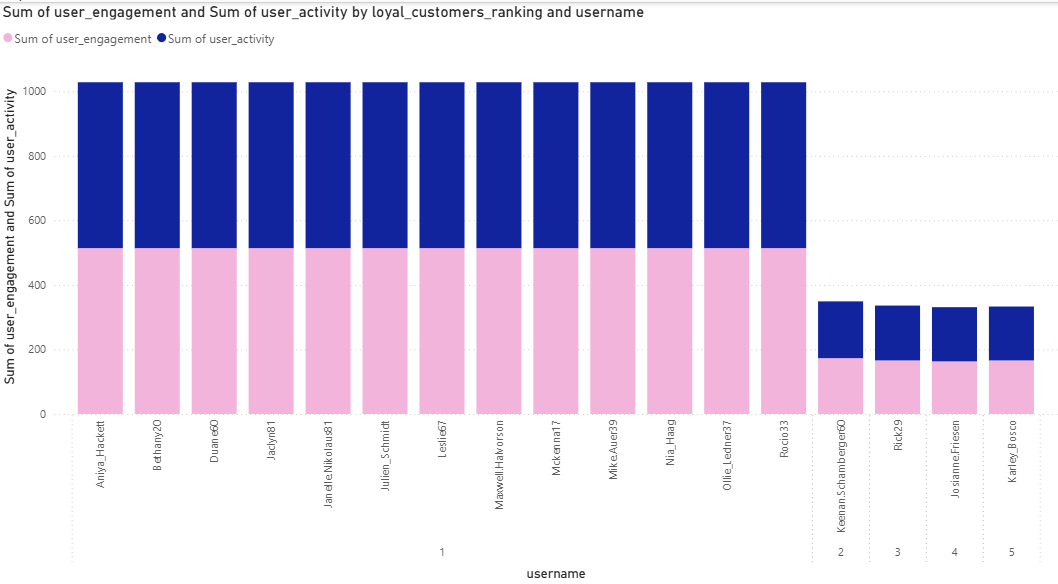
left join user\_photos d on a.id = d.user\_id)

select \*

from user\_table

where loyal\_customers\_ranking between 1 and 5;





**Insights**:

* The above table shows users with highest loyalty ranking (between 1 to 5) and are most active across all actions.
* The combined metric of user\_activity(likes and comments) recognizes both content creation and social interaction as likes are valuable for platform growth and comments are valuable for community building .

**Recommendations :**

* Implement a Loyalty Recognition Program  
  – Introduce tiered badges (e.g., “Top Contributor,” “Community Champion”) based on user activity levels to publicly recognize loyal users.
* Launch Exclusive Engagement Campaigns  
  – Offer early access to new features, limited-time filters, or platform upgrades to top-ranked users to incentivize continued interaction.
* Personalized Communication  
  – Send tailored appreciation emails or in-app messages acknowledging their contributions, with occasional discount codes or gift cards.

**2) For inactive users, what strategies would you recommend to re-engage them and**

**encourage them to start posting or engaging again?**

This query identifies users who have zero engagement on the platform,zero engagement as in  sum of all this ,Number of likes given by the user, number of comments made by the user, number of photos uploaded by the user i.e defined as engagement here is 0.

The query is:

select a.id,

a.username,

count(distinct b.user\_id)+count(distinct c.id)+count(distinct d.id) as engagement

from users a

left join likes b on a.id=b.user\_id

left join comments c on a.id=c.user\_id

left join photos d

on a.id=d.user\_id

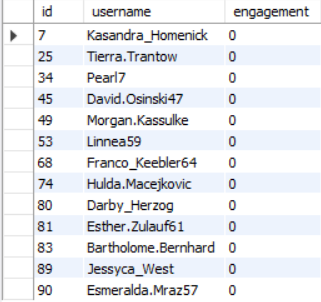
group by a.id,

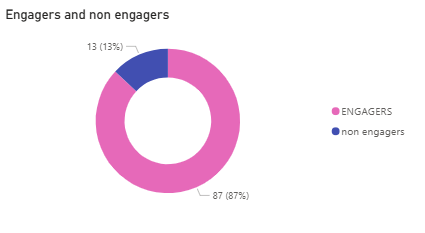
a.username

having count(distinct b.user\_id)+count(distinct c.id)+count(distinct d.id)=0;

**Insights**:

* + Below table shows all users with zero engagement:





* + These users signed up but have not participated in any meaningful way.
  + They might be lurkers, confused about how to engage, or simply lost interest.
  + Without action, they represent a risk for churn (leaving the platform altogether).

**Recommendations**:

* Highlight benefits of engagement: making friends, gaining followers, recognition.
* Remind inactive users of platform features with personalized messages like:
* “Hey [username], see what your friends posted today!”
* “Start sharing your moments — your followers are waiting.”
* Simplify Content Creation:
* Provide easy-to-use photo editing tools or templates inside the app.
* Allow quick posts (e.g., “one-tap” photo uploads).
* Integrate fun stickers, filters, and captions to lower the effort barrier.

**3)Which hashtags or content topics have the highest engagement rates? How**

**can this information guide content strategy and ad campaigns?**

with tag\_likes as(

select a.id as tag\_id,

a.tag\_name,

count(c.user\_id) as num\_likes

from tags a

left join photo\_tags b on a.id=b.tag\_id

left join likes c on b.photo\_id=c.photo\_id

group by tag\_id,tag\_name),

tag\_comments as(

select a.id as tag\_id,

a.tag\_name,

count(c.id) as num\_comments

from tags a left join photo\_tags b on a.id=b.tag\_id

left join comments c on b.photo\_id=c.photo\_id

group by tag\_id,tag\_name),

tag\_posts as(

select a.id as tag\_id,

a.tag\_name,

count(c.id) as num\_posts

from tags a

left join photo\_tags b on a.id=b.tag\_id

left join photos c on b.photo\_id=c.id

group by tag\_id,tag\_name)

select a.tag\_id,

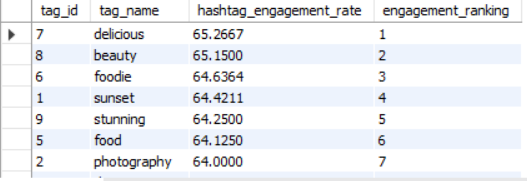
a.tag\_name,

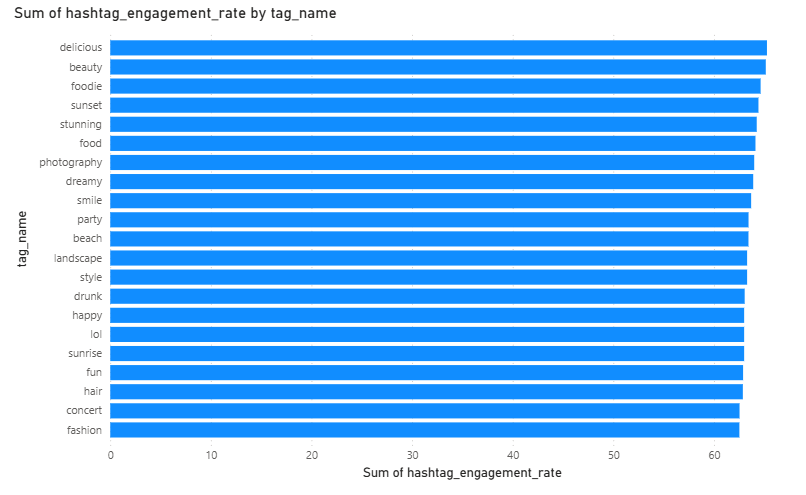
(num\_likes+num\_comments)/num\_posts as hashtag\_engagement\_rate ,

dense\_rank()over(order by (num\_likes+num\_comments)/num\_posts desc) as engagement\_ranking from tag\_likes a

join tag\_comments b on a.tag\_id=b.tag\_id

join tag\_posts c on a.tag\_id=c.tag\_id





**Insights**:

* Food-Related Hashtags Dominate
* delicious, foodie, and food all appear in the top 6.
* Users are highly engaged with food content, suggesting it’s both shareable and likable.
* Aesthetic and Visual Appeal Matter
* Tags like beauty, sunset, stunning, and photography also rank high.
* This suggests users engage with visually appealing content.

**Recommendations**:

* Leverage High-Engagement Tags in Explore/Trending sections.
  + - Promote posts with tags like #delicious, #beauty, and #sunset to new users to increase the initial stickiness.
* Paid Ad Targeting: Focus sponsored content on themes like food, beauty, and travel to tap into natural user interests.
* Encourage Users to Post in These Categories
* Recommend trending hashtags during post creation.
* Implement suggestions like: “People love #delicious – try it!”
* Content Challenges or Campaigns
* Weekly campaigns like “Best #SunsetShot” or “#FoodieFridays” can build habits and increase repeat usage.
* Hashtag-Based Feed Customization
* Provide personalized feeds around top engaging tags.

**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?**



Photo posting time:

with post\_posting as(

select

a.id as post\_id,

a.user\_id,

d.username,

a.created\_dat as post\_time,

count(b.user\_id)+count(c.id) as engagement

from photos a

left join likes b on a.id=b.photo\_id

left join comments c on a.id=c.photo\_id

join users d on d.id=a.user\_id

group by a.id,

a.user\_id,

a.created\_dat,

d.username)

select hour(post\_time) as post\_hr,

dayofweek(post\_time) as post\_day ,

count(post\_id) as num\_posts,

round(avg(engagement),0) as avg\_engagement

from post\_posting group by post\_hr,

post\_day;



User engaging with photos:

select dayname(a.created\_at) as day\_of\_week,

hour(a.created\_at) as hour\_of\_day,

count(distinct d.user\_id) as num\_comments,

count(distinct c.user\_id) as num\_likes,

count(distinct d.user\_id)+count(distinct c.user\_id) as engagement

from users a

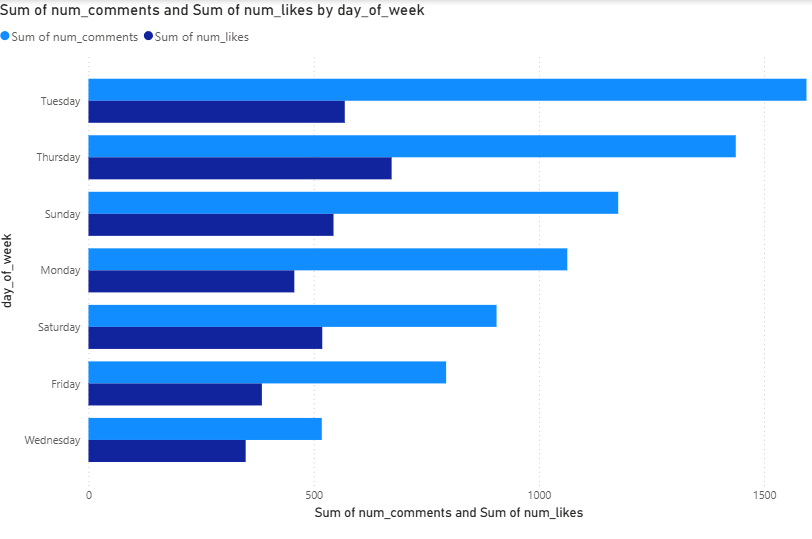
left join photos b on a.id=b.user\_id

left join likes c on b.id=c.photo\_id

left join comments d on b.id=d.photo\_id

group by dayname(a.created\_at),

hour(a.created\_at) ;



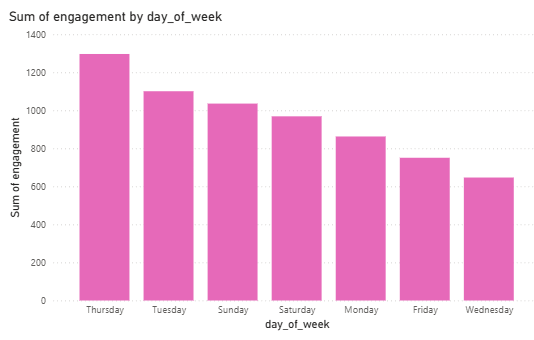
**For engagement throughout the week:**

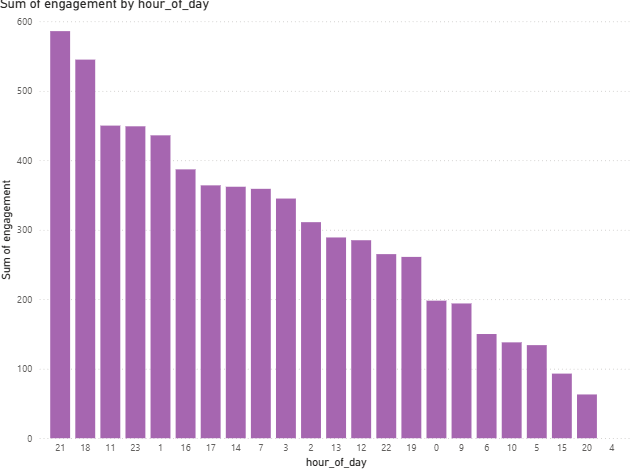
**Insights:**

1. Thursdays and Tuesdays emerged as the days with the highest engagement on posts, suggesting these are the most active days for user interaction.
2. Wednesdays showed the lowest engagement, indicating reduced user activity and interaction with content on this day.

**Recommendations:**

1. Encourage content scheduling on peak days (Thursday and Tuesday) to capitalize on higher user engagement and improve post reach.
2. Educate users and marketing teams about daily engagement trends to help them plan more effective posting strategies, avoiding low-performing days like Wednesday.





**For engagement throughout the day:**

**Insights:**

* The highest engagement is observed at 21:00 (586 engagements), followed by 18:00 (545).
* Engagement remains high between 11:00 to 23:00, indicating strong activity during late mornings through late evenings.
* Engagement significantly drops from midnight (0:00) to 10:00, with the lowest point at 4:00. Early morning and mid-afternoon are less optimal for user activity.

**Recommendations:**

* Optimize content delivery by posting between 18:00 and 23:00, with 21:00 as the prime time.This timing increases the likelihood of visibility and engagement.
* Avoid Low-Performing Hours
* Minimize or avoid scheduling posts between 2:00 and 10:00, especially 4:00, when engagement is almost nonexistent.

**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?**

with user\_followers as(

select followee\_id as id,

count(followee\_id) as no\_of\_followers

from follows

group by followee\_id)

,user\_likes as(

select user\_id,

count(\*) as num\_likes

from likes group by user\_id)

,user\_comments as(

select user\_id,

count(id) as num\_comments

from comments

group by user\_id)

,user\_photos as(

select user\_id,

count(id) as num\_photos

from photos

group by user\_id)

select d.username,

a.no\_of\_followers,

b.num\_likes,

c.num\_comments,

round((b.num\_likes+c.num\_comments)/num\_photos,2) as engagement\_rate,

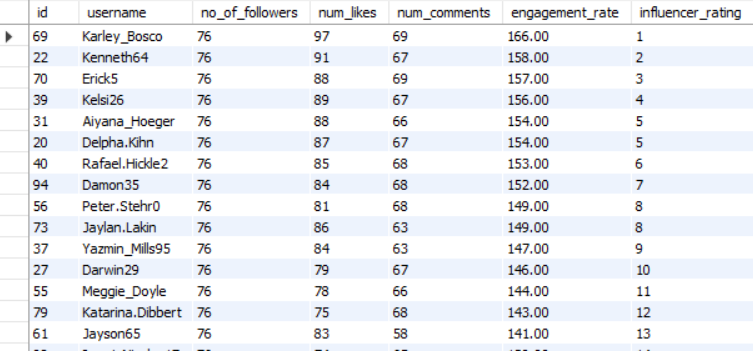
dense\_rank()over(order by b.num\_likes+c.num\_comments desc) as influencer\_rating from user\_followers

a join user\_likes b on a.id=b.user\_id

join user\_comments c on a.id=c.user\_id

join users d on a.id=d.id

join user\_photos e on a.id=e.user\_id;



**Insights**:

* Karley\_Bosco, Kenneth64, and Erick5 lead in engagement\_rate, making them prime influencer candidates.
* Engagement rates above 150 are considered extremely high for micro-influencers.
* Top users show a healthy mix of likes and comments, indicating content is not only liked but sparks interaction—a signal of genuine influence.
* All users have equal follower counts (76), but engagement rates vary from 144.00 to 166.00, indicating that content quality and audience fit matter more than raw audience size.

**Recommendations**:

* Target Top 5 Influencers for Paid Campaign Pilots
  + Focus on: Karley\_Bosco, Kenneth64, Erick5, Kelsi26, Aiyana\_Hoeger.
* Use Influencers to Drive Niche Campaigns by letting influencers co-create content tailored to their niche (e.g., tutorials, unboxings).
* Offer Long-Term Collaborations to Consistent Top Performers by converting high performers (e.g., Kenneth64, Erick5) into brand ambassadors.

**6)Based on user behavior and engagement data, how would you segment the user base**

**for targeted marketing campaigns or personalized Recommendations?**

with user\_likes as(

select user\_id,

count(\*) as num\_likes

from likes

group by user\_id)

,user\_comments as(

select user\_id,

count(id) as num\_comments

from comments

group by user\_id)

,user\_posts as(

select user\_id,

count(id) as num\_posts

from photos

group by user\_id),

users\_engagement as(

select a.id,

a.username,

coalesce(num\_likes,0)+coalesce(num\_comments,0)+coalesce(num\_posts,0) as user\_behaviour,

coalesce(num\_likes,0)+coalesce(num\_comments,0) as engagement

from users a

left join user\_likes b on

a.id=b.user\_id

left join user\_comments c on a.id=c.user\_id

left join user\_posts d on a.id=d.user\_id

)

select \*,

case

when engagement<=((select max(engagement)

from users\_engagement)-(select min(engagement) from users\_engagement))/3 then "low engagement"

when engagement>=2\*((select max(engagement) from users\_engagement)-(select min(engagement) from users\_engagement))/3 then "high engagement"

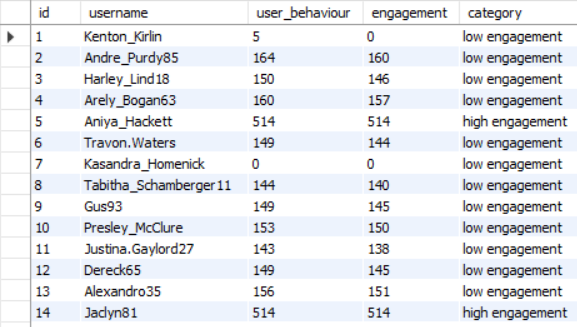
when engagement between ((select max(engagement) from users\_engagement)-(select min(engagement) from users\_engagement))/3 and

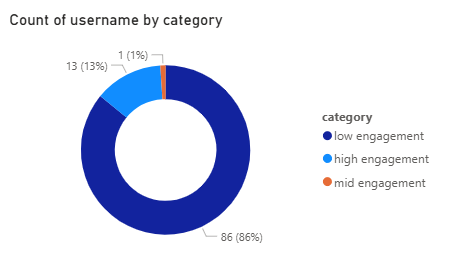
2\*((select max(engagement) from users\_engagement)-(select min(engagement) from users\_engagement))/3

then "mid engagement"

end as category

from users\_engagement;



****

**Insights**:

* Over 85% of users fall into the low engagement segment despite moderate user behavior scores (likes, comments, posts).
* Aniya\_Hackett and Jadyn81 have exceptionally high engagement and user activity (514).
* Users like Kasandra\_Homenick (0 across the board) and Kenton\_Kirlin (very low behavior and engagement) show signs of churn.

**Recommendations**:

* Launch re-engagement campaigns targeting low-engagement users using personalized feeds to surface relevant content.
* Involve top performers like Aniya\_Hackett, Jadyn81 in beta testing or early access features
* For users posting often but receiving low interaction, provide content creation tips (e.g., posting time, hashtags)
* For zero or near-zero activity users, send personalized reactivation emails/messages.

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

Key Metrics to Track for Effectiveness

* Click-Through Rate (CTR) = (Clicks / Impressions) \* 100

CTR gauges how many clicked on your ad out of the people who viewed it. The better the CTR, the more effective your ad creatives and targeting are at catching attention. It aids in gauging the relevance and attractiveness of your ad to the audience.

* Conversion Rate (CVR) = (Conversions / Clicks) \* 100

CVR informs you about the conversion rate of your ad traffic into significant actions such as purchase, sign-up, or downloading. It indicates the quality of your traffic as well as the landing page or offer effectiveness.

* Cost Per Acquisition (CPA) = Total Spend / Conversions

CPA indicates how much you're spending to get one customer or desired action. A lower CPA means your campaign is cost-effective. It's important for budgeting and ROI analysis.

* Return on Ad Spend (ROAS) = Revenue / Total Spend

ROAS is a measure of profitability. It shows you how much revenue you are making per ad dollar spent. If your ROAS is greater than 1, then you are making more than you are spending, which is the end objective of any campaign.

* Impressions and Reach

These show how many users viewed the ad and how widely it reached. They assist in comprehending the visibility of the campaign and whether you're reaching your targeted audience size and scope.

* Future Optimization Strategies  
  Reallocate Budget to Top-Performing Segments  
  Determine which audience segments, devices, or locations are producing higher CTR, CVR, and ROAS, and allocate more budget to them. This maximizes the use of funds to be focused in profitable areas.
* Pause or Optimize Underperforming Campaigns  
  Find ads with high spend but low conversion or engagement. Pause them or test new creatives, copy, or targeting to enhance performance without burning budget.
* Retarget Users Who Clicked but Failed to Convert  
  Create retargeting campaigns for users who were interested but failed to perform the desired action. These are already warm leads, and with a soft reminder or an improved offer, they will easily convert.
* Optimize Ad Creatives and Messaging Through A/B Testing  
  Test various iterations of your ads visuals, headlines, calls-to-action — to determine which combinations perform better. Ongoing testing serves to reveal winning ingredients and minimize guesswork

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

In order to successfully identify those users that might be good brand ambassadors or advocates for Instagram campaigns, we can scrutinize a few important features of their activity and profile information:

* Engagement Rate: Check for users who always get high levels of engagement on their posts—likes, comments, shares, saves. A high level of engagement shows that their followers are highly interested and responsive to their post, hence influential in propagating messages effectively.
* Relevant Hashtag Usage: Engage users who habitually use hashtags that fit Instagram's promotional theme or goal. This indicates that such users are already interested or engaged in what the initiatives are about, making them more likely to support and promote the events genuinely.
* Follower Number: Target users with a moderate to large following. Although very large followings are worth the investment, mid-influencers tend to have more loyal audiences and can generate substantial engagement with more genuine interactions.

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

Approach :

1. Analyze Followers and Following Patterns

* Identify users with high follower counts to assess their potential influence and reach.
* Examine the number of users each individual follows to understand their content consumption behavior.
* Calculate the follower-to-following ratio to differentiate between influencers and regular users.
* Highlight mutual connections (bidirectional follows) to detect community engagement or close user groups.

2. Study User Posting Activity

* Count the total number of posts each user has made to gauge content creation levels.
* Categorize users based on posting frequency:
  + High Activity: Frequent posters who may drive content flow.
  + Moderate Activity: Occasionally active users.
  + Low Activity: Rarely post content.
  + Non-Posters: Users who’ve never posted—possible churn risks.
* Analyze the average interval between posts to assess consistency in user activity.

3. Evaluate Engagement Metrics (Likes and Comments)

* Calculate total likes and comments received per user to understand content reception.
* Derive engagement rate by dividing likes and comments by total followers to measure content effectiveness.
* Identify active engagers—users who frequently like or comment—to nurture and involve them more in community campaigns.
* Detect “ghost followers” who never interact with posts and might be passive or bot accounts.

4. Explore Hashtag Usage and Trends

* Analyze which hashtags are most frequently used across the platform.
* Measure the average engagement (likes/comments) per hashtag to identify those that lead to better visibility and interaction.
* Recognize users who frequently use trending or diverse hashtags, suggesting an understanding of platform dynamics.
* Compare engagement metrics between tagged and non-tagged posts to validate the importance of hashtag usage.

**Insights** That Can Be Derived

1. High follower users may act as influencers and help amplify reach organically.
2. Users who post regularly are more likely to retain attention and gain engagement.
3. A significant portion of users may not post at all—targeting them can unlock dormant engagement potential.
4. Certain hashtags might be driving high engagement; promoting these could benefit the broader user base.
5. Personalized engagement (comments) often signals deeper connection than passive likes—content formats generating comments should be scaled.
6. Identifying low-engagement zones (inactive users, unpopular hashtags) can help redirect strategy and resources.

**Recommendations** to Improve Engagement, Retention, and Acquisition

1. Personalized Content and Feed Optimization

* Recommend posts based on individual user preferences, recent activity, and similar profiles.
* Use machine learning to highlight engaging content formats (e.g., carousel posts, reels, story types).

2. Reward-Based Posting and Interaction

* Introduce gamified rewards for users who hit engagement milestones like “100 likes” or “10 shares”.
* Offer premium filters, stickers, or badge unlocks as incentives for frequent posters or commenters.

3. Community and Influencer Engagement

* Partner with high-engagement users to promote challenges, reels, or campaigns.
* Run exclusive creator programs to feature their content and boost follower count organically.

4. Referral and Growth Incentives

* Create a referral program where users are rewarded for inviting new users.
* Offer limited-time perks to both inviter and new user to boost sign-ups and onboarding.

5. Re-engagement of Dormant Users

* Send personalized push notifications to users who haven’t posted or engaged in a while.
* Highlight what they’re missing (e.g., “5 friends posted today, come see what’s new!”).

6. Hashtag Promotion and Education

* Educate users on the impact of hashtags via in-app tutorials or nudges.
* Automatically suggest high-performing hashtags while posting to improve discoverability.

**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?**

Update User\_Interactions

Set Engagement\_Type=”Heart”

Where Engagement\_Type=”Like” ;