**Instagram User Analytics**

**Project Description:** This project isabout user analysis. I was asked to perform user analysis for Instagram and provide insights to product manager on the questions asked by the management team. First it was required to create the database using dataset provided. Then it was required to perform analysis using SQL to answer the below questions:

**A) Marketing:**

1. Rewarding Most Loyal Users: Find the 5 oldest users of the Instagram from the database provided
2. Remind Inactive Users to Start Posting: Find the users who have never posted a single photo on Instagram
3. Declaring Contest Winner: Identify the winner of the contest and provide their details to the team
4. Hashtag Researching: Identify and suggest the top 5 most commonly used hashtags on the platform
5. Launch AD Campaign: What day of the week do most users register on? Provide insights on when to schedule an ad campaign

**B) Investor Metrics:**

1. User Engagement: Provide how many times does average user posts on Instagram. Also, provide the total number of photos on Instagram/total number of users
2. Bots & Fake Accounts: Provide data on users (bots) who have liked every single photo on the site (since any normal user would not be able to do this).

**Approach:** First Igone through all the tables to know all the columns present in each table. Then I saw all the questions and thought of which tables could be used in each question. Then I wrote the sql queries for each question.

**Tech-Stack Used:** The software used for the project is MySQL workbench 8.0 CE. It is used to run the sql queries and get answers of each question.

**Insights:**

**A) Marketing:**

1. Rewarding Most Loyal Users:

In this, I was required to find 5 oldest users of Instagram.

Sql query:

SELECT \*

FROM (

SELECT

id, username, created\_at,

RANK() OVER(order by created\_at) as joining\_rank

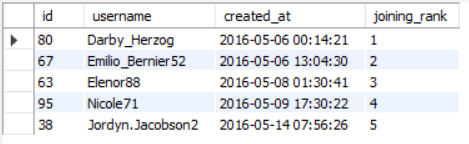
FROM

users) as ranking

WHERE

joining\_rank < 6;

Output:



1. Remind Inactive Users to Start Posting:

In this, I was required to find users who have never posted a single photo on Instagram.

Sql query:

SELECT

id, username

FROM

users

WHERE

id NOT IN (

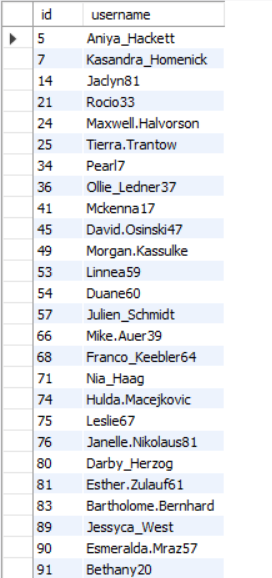
SELECT

user\_id

FROM

photos);

Output:



1. Declaring Contest Winner:

In this, I was required to find the details of the user who gets the most likes on a single photo.

Sql query:

CREATE VIEW likes\_view AS

(SELECT

photo\_id, COUNT(photo\_id) AS no\_of\_likes

FROM

likes

GROUP BY photo\_id);

SELECT

photo\_id, user\_id, username, no\_of\_likes

FROM

users INNER JOIN

(SELECT

photo\_id, user\_id, no\_of\_likes

FROM

likes\_view INNER JOIN

photos ON likes\_view.photo\_id = photos.id

WHERE

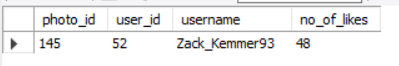
no\_of\_likes = (SELECT

MAX(no\_of\_likes)

FROM

likes\_view)) AS inner\_query ON users.id = inner\_query.user\_id;

Output:



1. Hashtag Researching: Identify and suggest the top 5 most commonly used hashtags on the platform

Sql query:

CREATE VIEW hashtags\_view AS

(SELECT

tag\_id, COUNT(tag\_id) AS used\_no

FROM

photo\_tags

GROUP BY tag\_id);

SELECT \*

FROM (

SELECT

tag\_id, tag\_name, used\_no,

RANK() OVER(order by used\_no desc) as hashtag\_rank

FROM

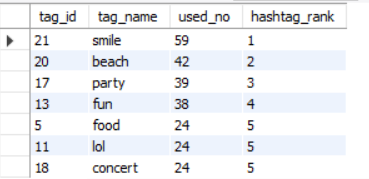
hashtags\_view inner join

tags on hashtags\_view.tag\_id = tags.id) as ranking

WHERE

hashtag\_rank < 6;

Output:



1. Launch AD Campaign: What day of the week do most users register on? Provide insights on when to schedule an ad campaign

Sql query:

CREATE VIEW users\_registered\_view AS

(SELECT

DAYOFWEEK(created\_at) as Day\_of\_week,

COUNT(DAYOFWEEK(created\_at)) AS Users\_registered

FROM

users

GROUP BY DAYOFWEEK(created\_at)

ORDER BY Day\_of\_week);

SELECT

Day\_of\_week, Users\_registered AS Most\_users\_registerd

FROM

users\_registered\_view

WHERE

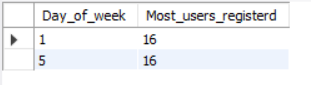
Users\_registered = (SELECT

MAX(Users\_registered)

FROM

users\_registered\_view);

Output:



**B) Investor Metrics:**

1. User Engagement: Provide how many times does average user posts on Instagram. Also, provide the total number of photos on Instagram/total number of users

Sql query:

CREATE VIEW posts\_view AS

(SELECT

user\_id, COUNT(id) AS no\_of\_posts

FROM

photos

GROUP BY user\_id

ORDER BY user\_id);

SELECT

id, username, no\_of\_posts

FROM

users

LEFT JOIN

posts\_view ON users.id = posts\_view.user\_id;

CREATE VIEW posts\_per\_user AS

(SELECT

id, username, no\_of\_posts

FROM

users

LEFT JOIN

posts\_view ON users.id = posts\_view.user\_id);

SELECT \* FROM posts\_per\_user;

SELECT

COUNT(id) AS no\_of\_users,

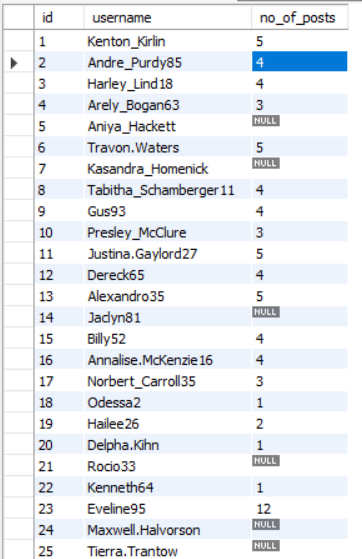
SUM(no\_of\_posts) AS total\_no\_of\_posts,

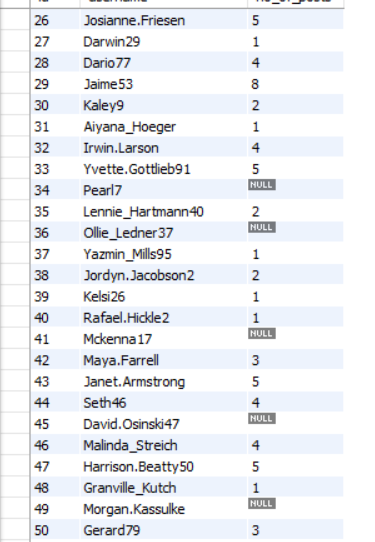
SUM(no\_of\_posts) / COUNT(id) AS `total number of photos on Instagram/total number of users`

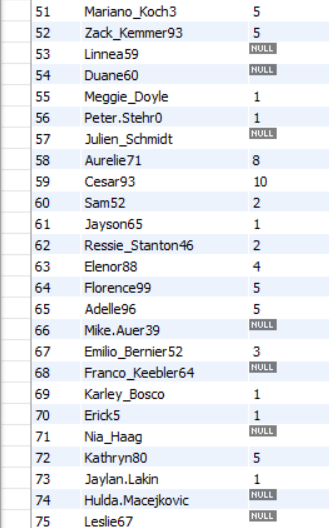
FROM

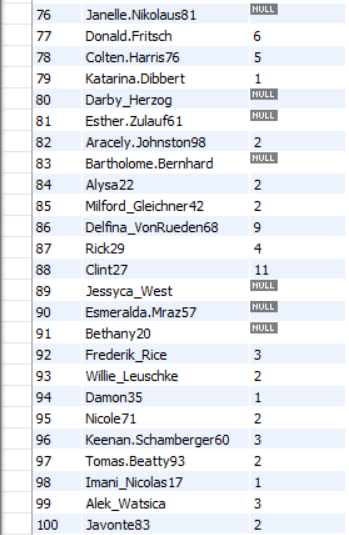
posts\_per\_user;

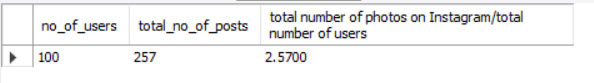
Output:











1. Bots & Fake Accounts: Provide data on users (bots) who have liked every single photo on the site (since any normal user would not be able to do this)

Sql query:

SELECT

user\_id, username

FROM

(SELECT

user\_id, username, COUNT(photo\_id) AS posts\_liked

FROM

likes

INNER JOIN users ON likes.user\_id = users.id

GROUP BY user\_id) AS likes\_no

WHERE

posts\_liked = (SELECT

COUNT(DISTINCT id) AS total\_posts

FROM

photos);

Output:



**Results:**

**A) Marketing:**

1. Rewarding Most Loyal Users:

Darby\_Herzog, Emilio\_Bernier52, Elenor88, Nicole71, Jordyn.Jacobson2 with user ids 80, 67, 63, 95, 38 respectively are five oldest users of Instagram.

1. Remind Inactive Users to Start Posting:

Aniya\_Hackett, Kasandra\_Homenick, Jaclyn81, Rocio33, Maxwell.Halvorson, Tierra.Trantow, Pearl7, Ollie\_Ledner37, Mckenna17

David.Osinski47, Morgan.Kassulke, Linnea59, Duane60, Julien\_Schmidt, Mike.Auer39, Franco\_Keebler64, Nia\_Haag, Hulda.Macejkovic, Leslie67, Janelle.Nikolaus81, Darby\_Herzog, Esther.Zulauf61, Bartholome.Bernhard, Jessyca\_West, Esmeralda.Mraz57, Bethany20 are the users who have never posted a single photo on Instagram

1. Declaring Contest Winner:

Zack\_Kemmer93 with user id 52 is winner of the contest by getting most likes on a single photo.

1. Hashtag Researching:

smile, beach, party, fun, food are the five most used hashtags on the platform. Also, lol and concert are used same number of times as food.

1. Launch AD Campaign:

Most users register on day 1 and day 5 of week i.e., Sunday and Thursday. According to me, ad campaign could be launched on the day when most number of likes are created which mean on the day when posts engagement are highest.

**B) Investor Metrics:**

1. User Engagement:

Total number of photos on Instagram/total number of users = 2.57

1. Bots & Fake Accounts:

Aniya\_Hackett, Jaclyn81, Rocio33, Maxwell.Halvorson, Ollie\_Ledner37, Mckenna17, Duane60, Julien\_Schmidt, Mike.Auer39, Nia\_Haag, Leslie67

Janelle.Nikolaus81, Bethany20 have liked each and every photo on platform.