**Instagram User Analytics**

**Project Description:** This project helps to analyse and visualize raw data using different database tools to get useful insights which will be beneficial for the business.

**Project Approach:** SQL was used to execute the project in which database was created using the raw data. Then the extraction was done through below queries to get the required outputs.

1. **Marketing**
2. **Rewarding Most Loyal Users:** Find the 5 oldest users of the Instagram from the database provided

select id,

username,

DATE(created\_at) from users

group by id,username, Date(created\_at)

order by Date(created\_at)

limit 5;

**Explanation of above query:** Id, username & Date (extracted from created\_at column) are selected from table users, grouped as per id and username and finally arranged in ascending order of date so that the oldest dates can be found. Limit 5 is used to get 5 oldest users.

1. **Remind Inactive Users to Start Posting**:  Find the users who have never posted a single photo on Instagram

select U.\* from users U

where U.id not in

(select P.user\_id from photos P);

**Explanation of above query:** First, all the details from table users were selected in line 1. Query 2 was executed to find the user id from table photos. Then, both the query was attached using a condition where id in users table is not present in query 2.

1. **Declaring Contest Winner:** Identify the winner (most likes on a single photo) of the contest and provide their details to the team

SELECT U.username,L.photo\_id,PL.likes\_count

from likes L

join

(select photo\_id,count(\*) as likes\_count from likes

group by photo\_id) as PL

join users U

ON U.id = L.user\_id

order by likes\_count desc

limit 1;

**Explanation of above query:** Details of Photo\_id and count of all records (named as like\_count), grouped according to photo\_id are extracted and named the table as PL. Then table likes and PL are joined to extract username, photo\_id and likes\_count.

Limit 1 is used to get the winner as the most likes on a single photo.

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

select PT.tag\_id,T.tag\_name,count(\*) as tag\_count from tags T

join photo\_tags PT

on T.id = PT.tag\_id

group by PT.tag\_id,T.tag\_name

order by tag\_count desc

LIMIT 5;

**Explanation of above query:** Tag\_id, tag\_name and count of all records (named as tag\_count) are extracted from table tags. Then, joined with the table photo tags as per the group of tag\_id and tag\_name. The output was ordered in the descending order of tag count.

Limit 5 is used to get top 5 most commonly used hashtags.

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

select DAYNAME(created\_at) as week\_day,

count(\*) as user\_count from users

group by week\_day

order by user\_count desc;

**Explanation of above query:** Dayname was extracted from created\_at column of table users and named as week day. Then, dayname & count of all records are extracted, grouped by week day and arranged in descending order of count to get the final output.

**Insights**: Most users register on Thursday and Sunday. So, ad campaigns should be schedule on these days to increase the business.

1. **Investor Metrics**
2. **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

select count(distinct user\_id) as total\_users,

count(image\_url) as total\_photos,

count(\*)/count(distinct user\_id) as avg\_user\_posts

from photos;

**Explanation of above query:** count of users by distinct user id as total number of users and count of image url as total number of photos are calculated. Then, total records are divided by count of users to get the average user posts.

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).

select P.user\_id,

count(P.user\_id) as user\_count,

count(L.user\_id) as likes\_count from photos P

join likes L

on P.user\_id = L.user\_id

group by P.user\_id;

**Explanation of above query:** Count of user id as user count from table photos and count of likes as likes count from table likes are calculated by joining two tables’ photos and likes. Then grouped according to user id to get the required output.

**Tech-Stack Used:** MySQL workbench is used for project execution. The purpose is to extract the data from database.

**Final Insights:**

* Total number of users registered.
* Loyal customers.
* The active and non-active users.
* Alerts to be sent to inactive users.
* The fake and genuine number of accounts present.
* Sunday and Thursday are the days when most users logged in to Instagram.
* Users who have registered but not posted a single photo.
* Most commonly used hashtags.