Instagram user analytics project

PROJECT DESCRIPTION-

Instagram, as we all must be knowing that is one of the most used social sites in today's world. It's really addicting right? You can share your favorite photo or video. It allows users to upload media that can be edited with various inbuilt filters and hashtags and geographical tagging. You also get options to share with your preapproved followers or publicly. It is really, interesting to think that where does all the data gets stored when we like or comment on someone's picture. To understand all these . I created a small project that replicates some of the features of Instagram. As we proceed above, I created a database by the name Instagram the database consists of total 7 tables.

1.USER TABLE

2.PHOTOS TABLE

3.COMMENTS TABLE

4.LIKES TABLE

5.FOLLOWS TABLE.

6.TAGS TABLE

7.PHOTO-TAGS TABLE

• DETAILED REPORT

- A)

- MARKETING- The team wants to launch some campaigns, and they need some help with following

- questions-1.

- Rewarding the most loyal users: It shows the people who have been using the platform for the longest time.

- Task

- Find the 5 oldest users of the Instagram from the database provided by the team. Stack used .

- It tell your database that you want to select data. FROM users tells the database to select data from the user table.(*) tells the database that you want to see all columns in this table. ORDER BY- after this expression, simply specify a column on which the data will be sorted. LIMIT n - returns the first n rows from the result. This is much more efficient than returning all the data from the database.

```sql
1 • SELECT `users`.`id`,
2        `users`.`username`,
3        `users`.`created_at`
4  FROM `instagram`.`users`
5  order by created_at ASC
6  limit 5;
```

Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content: | Fetch rows:

id	username	created_at
80	Darby_Herzog	2016-05-06 00:14:21
67	Emilio_Bernier52	2016-05-06 13:04:30
63	Elenor88	2016-05-08 01:30:41
95	Nicole71	2016-05-09 17:30:22
38	Jordyn.Jacobson2	2016-05-14 07:56:26
NULL	NULL	NULL

- 2.

- Inactive users on instagram Task - Find the users who have never posted a single photo on Instagram.
- works in the following way: it returns all rows from the left table (the first table in the query) plus all matching rows from the right table (the second table in the query).

NOTE: ALL OUT RESULT WERE NOT COMING IN SCREENSHOT BUT QUERY IS CORRECT


```sql
1  SELECT username FROM instagram.users  LEFT JOIN instagram.photos ON users.id = photos.user_id WHERE photos.id IS NULL
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

username
Aniya_Hackett
Kasandra_Homenick
Jaclyn81
Rocio33
Maxwell.Halvorson
Tierra.Trantow
Pearl7
Ollie_Ledner37

3. <u>Declaring contest winner</u>

- Task: Identify the winner of the contest and provide their details to the team. To do this task we need to find the most popular photo with most likes and users who created it. SQL QUERY

INNER JOIN (or JOIN)

only shows those rows from the two table somewhere there is a match between the columns. In other words, you can only see those pieces of equipment which have a room assigned and vice versa

```
1   use instagram;
2 • SELECT username,photos.id,photos.image_url,COUNT(*) AS total FROM photos
3   INNER JOIN likes ON likes.photo_id = photos.id INNER JOIN users ON photos.user_id = users.id GROUP BY photos.id ORDER BY total DESC LIMIT 1;
4
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

username	id	image_url	total
Zack_Kemmer93	145	https://jarret.name	48

4. HASHTAG RESEARCHING

The Hashtag helps the user to reach a wide range of people. It is used to draw attention, organize, promote the connect.

Task- to identify the top 5 most commonly used hashtags on instagram.

AS- The new name is just an alias, which means it's temporary and doesn't change the actual column name in the database. It only influences the way the column is shown in the result of the specific query. This technique is often used when there are a few columns with the same name coming from different tables.
Normally, when SQL displays columns in the result, there is no information about the table that a specific column is part of.

```
1   use instagram;
2   SELECT tags.tag_name,COUNT(*) AS total FROM photo_tags JOIN tags ON photo_tags.tag_id = tags.id GROUP BY tags.id ORDER BY total DESC LIMIT 5;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

tag_name	total
smile	59
beach	42
party	39
fun	38
concert	24

5. LAUNCH AD CAMPAIGN

- TASK-To find out the day of week when most users register's on Instagram.

- The query result shows two days of the week when the users register mostly According to me the most suitable day to launch the ad campaign would be Sunday because users mostly have leisure time on Sunday and more interaction would happen on Sunday.

Limit to 1000 rows ▾

```
1    SELECT DAYNAME(created_at) AS day,COUNT(*) AS total FROM instagram.users GROUP BY day ORDER BY total DESC LIMIT 2;
```

Result Grid | 🔢 ⇄ Filter Rows: | Export: 🖫 | Wrap Cell Content: 🛱 | Fetch rows:

day	total
Thursday	16
Sunday	16

B. <u>INVESTOR MATRICS</u>

6. USER ENGAGEMENT- Investors want to know that instagram is not becoming redundant like facebook, so they want to check the frequency of how much the users are engaging on the platform.

Task –To provide how many times an average user post on instagram.

Limit to 1000 rows

```
1    SELECT (SELECT Count(*) FROM instagram.photos) / (SELECT Count(*) FROM instagram.users) AS avg;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

avg
2.5700

- 2. BOTS& FAKE ACCOUNTS- It is reported that there are lot of bots and fake accounts on the platform. The investors wants to know if there are fake and dummy accounts.

- TASK-To Provide data on users(bots) who have liked every single photo on the site(normal user would not be able to do this).

```sql
1 •    use instagram;
2
3 •    SELECT username,Count(*) AS num_likes FROM instagram.users
4        INNER JOIN likes ON users.id = likes.user_id GROUP BY likes.user_id HAVING num_likes = (SELECT Count(*)FROM instagram.photos);
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

username	num_likes
Aniya_Hackett	257
Jaclyn81	257
Rocio33	257
Maxwell.Halvorson	257
Ollie_Ledner37	257
Mckenna17	257
Duane60	257
Julien_Schmidt	257
Mike.Auer39	257
Nia_Haag	257
Leslie67	257
Janelle.Nikolaus81	257
Bethany20	257

OVERALL RESULT

While doing this project I found out so many important terms of MYSQL that helps in solving complex problems irrespective of how large the database is. I got to learn about using sql and where abouts of sql workbench and i believe it will help me in future a lot. I have provided the solutions to every questions asked and i believe they are correct to the best of my knowledge and it solves all the query.

THANKYOU.