



Instagram User Analytics

SQL Fundamentals

By Rajat Panwan



Project Description

In this project, MySQL commands are used to answer the questions asked by our investors and marketing team. The name of the database is “ig_clone” where I performed data manipulation on various tables such as, comments, follows, likes, photo_tags, photos, tags and users. Joining data from multiple tables using Inner Join, Left Join, and Right Join is also performed.

Approach

In this project I carefully read and understand the requirements and objective of the project. Then I go through the entire tables of the database to know tables attributes. In tables I use appropriate clauses such as Select, From, Order By, Group By and Where to extract the required information from the database. Once I had formulated the final query, I tested it thoroughly to ensure that it was correctly returning the desired results.

Tech-Stack Used

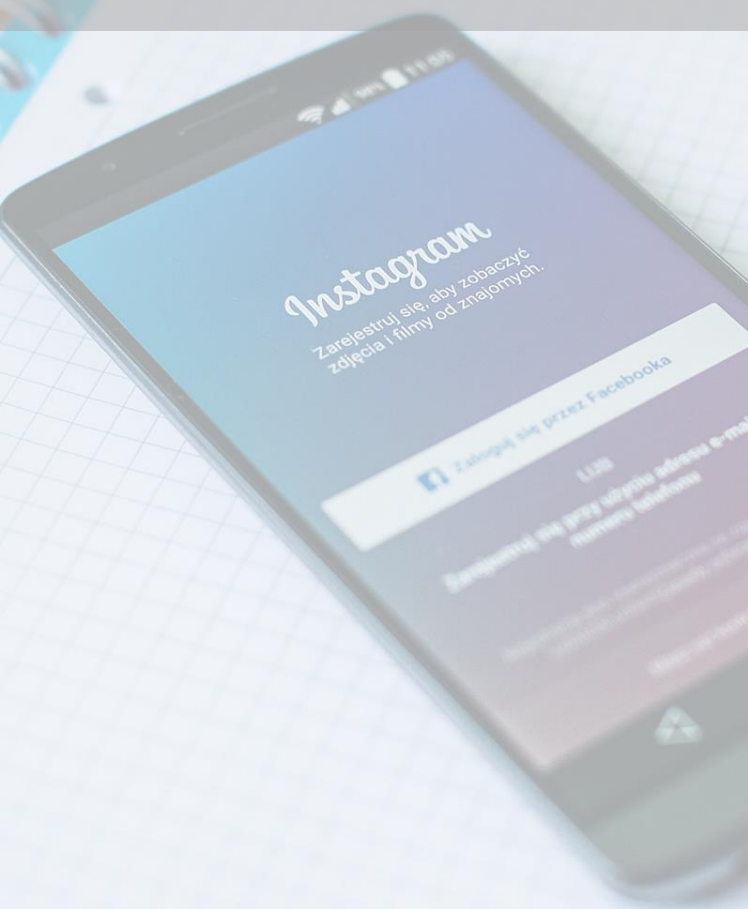
A database management system MySQL 8.0 is used to handle, store and modify and delete data and also store data in an organized way. In this process MySQL Workbench is used which comes with MySQL.



Insights

Here are some insights and knowledge that I gained while working on Instagram User Analytics project such as understating of the SQL language and how to use it to retrieve and manipulate data in a database. Develop an ability to design and execute complex queries using a range of SQL clauses, functions and operators. Skills in data analysis and problem solving as the process of creating an SQL query often involves identifying patterns in the data.

Result



Result of the queries I write is on the next pages.



Marketing

1. **Rewarding Most Loyal Users:** People who have been using the platform for the longest time.

Your Task: Find the 5 oldest users of the Instagram from the database provided.

Query:

```
1 • SELECT *
2 FROM users
3 ORDER BY created_at
4 LIMIT 5;
```

Result:

	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.Jacobs...	2016-05-14 07:56:26
●	NULL	NULL	NULL

Marketing

2. Remind Inactive Users to Start Posting: By sending them promotional emails to post their 1st photo.


Your Task: Find the users who have never posted a single photo on Instagram.


Query:


```
1 • SELECT *
2 FROM users
3 LEFT JOIN photos
4 ON users.id = photos.user_id
5 WHERE photos.user_id IS NULL;
```

Result:


Result Grid

 Filter Rows:

Export: 

Wrap Cell Content: 

	id	username	created_at	id	image_url	user_id	created_at
▶	5	Aniya_Hackett	2016-12-07 01:04:39	NULL	NULL	NULL	NULL
	7	Kasandra_Ho...	2016-12-12 06:50:08	NULL	NULL	NULL	NULL
	14	Jadyn81	2017-02-06 23:29:16	NULL	NULL	NULL	NULL
	21	Rocio33	2017-01-23 11:51:15	NULL	NULL	NULL	NULL
	24	Maxwell.Halvo...	2017-04-18 02:32:44	NULL	NULL	NULL	NULL
	25	Tierra.Trantow	2016-10-03 12:49:21	NULL	NULL	NULL	NULL
	34	Pearl7	2016-07-08 21:42:01	NULL	NULL	NULL	NULL
	36	Ollie_Ledner37	2016-08-04 15:42:20	NULL	NULL	NULL	NULL
	41	Mckenna17	2016-07-17 17:25:45	NULL	NULL	NULL	NULL

Result 1 

Marketing

3. Declaring Contest Winner:

The team started a contest and the user who gets the most likes on a single photo will win the contest now they wish to declare the winner.

Your Task: Identify the winner of the contest and provide their details to the team

Query:

```
1 • SELECT
2     *
3 FROM
4     (SELECT
5         *
6     FROM
7         (SELECT
8             MAX(likes_count) AS max_likes, photo_id
9         FROM
10            (SELECT
11                COUNT(photo_id) likes_count, photo_id
12            FROM
13                likes
14            GROUP BY photo_id
15            ORDER BY likes_count DESC) AS photo_likes) AS max_liked_photo
16        INNER JOIN photos ON photos.id = max_liked_photo.photo_id) AS most_liked_user_photo
17        INNER JOIN
18        users ON users.id = most_liked_user_photo.user_id;
```

Result:

Result Grid									
Filter Rows:									
Export: Wrap Cell Content:									
	max_likes	photo_id	id	image_url	user_id	created_at	id	username	created_at
▶	48	145	145	https://jarret.name	52	2022-11-28 18:37:28	52	Zack_Kemmer93	2017-01-01 05:58:22

Marketing

4. Hashtag Researching: A partner brand wants to know, which hashtags to use in the post to reach the most people on the platform.

Your Task: Identify and suggest the top 5 most commonly used hashtags on the platform

Query:

```
1 • SELECT
2     most_used_tags.most_used_tag, tags.tag_name
3 FROM
4     (SELECT
5         COUNT(tag_id) most_used_tag, tag_id
6     FROM
7         photo_tags
8     GROUP BY tag_id
9     ORDER BY most_used_tag DESC) AS most_used_tags
10    LEFT JOIN
11    tags ON tags.id = most_used_tags.tag_id
12    LIMIT 5;
```

Result:

Result Grid			Filter Rows:	Export:
	most_used_tag	tag_name		
▶	59	smile		
	42	beach		
	39	party		
	38	fun		
	24	food		

Result 1 ×

Marketing

5. Launch AD Campaign: The team wants to know, which day would be the best day to launch ADs.

Your Task: What day of the week do most users register on? Provide insights on when to schedule an ad campaign

Query:

```
1 • SELECT
2     COUNT(weekdays) weekday_counts, weekdays
3 FROM
4     (SELECT
5         DAYNAME(created_at) weekdays
6     FROM
7         users) AS week_table
8 GROUP BY weekdays
9 ORDER BY weekday_counts DESC;
```

Result:

Result Grid			Filter Rows:	Export:	Wrap Cell Content:
	weekday_counts	weekdays			
▶	16	Thursday			
	16	Sunday			
	15	Friday			
	14	Tuesday			
	14	Monday			
	13	Wednesday			
	12	Saturday			

Result 1 x

Investor Metrics

1. User Engagement: Are users still as active and post on Instagram or they are making fewer posts

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

Query:

```
11 • SELECT
12 ((SELECT
13     COUNT(id)
14 FROM
15     photos) / (SELECT
16     COUNT(id)
17 FROM
18     users)) AS divide;
19
20
```

Result:

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
	Average_Post_per_User			
▶	3.4730			

Result 1 ×

Investor Metrics

2. Bots & Fake Accounts: The investors want to know if the platform is crowded with fake and dummy accounts

Your Task: 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).

Query:

```
1 • SELECT
2     user_id, COUNT(user_id) AS like_count
3 FROM
4     likes
5 GROUP BY user_id
6 HAVING COUNT(user_id) = (SELECT
7     COUNT(DISTINCT (photo_id)) AS distinct_photo_id
8 FROM
9     likes);
```

Result:

	user_id	like_count
▶	5	257
	14	257
	21	257
	24	257
	36	257
	41	257
	54	257
	57	257
	66	257
	71	257
	75	257

Result 1 x

Thank You

You can connect me on:
rajatpawan@gmail.com