Instagram User Analytics

Project - 2

TITLE- Instagram User Analytics

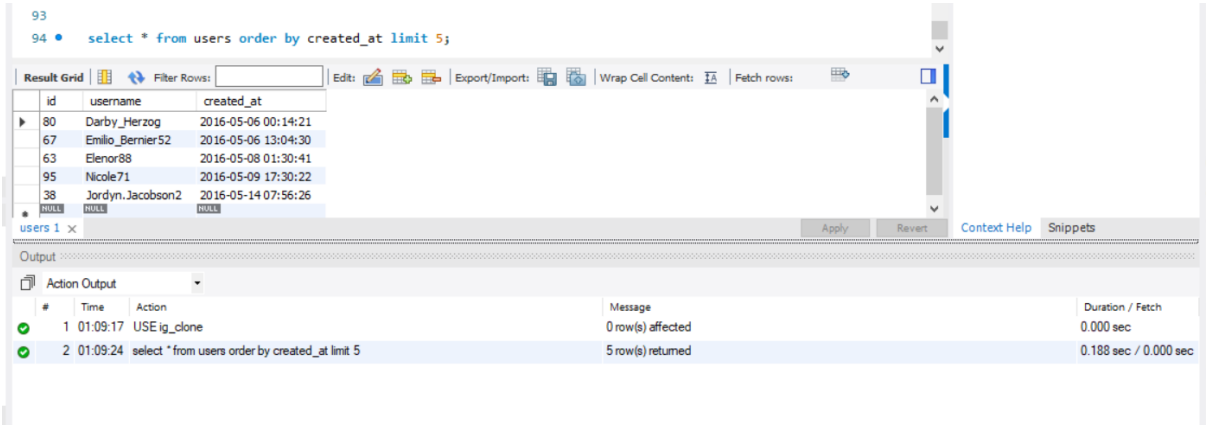
Project Overview:

The purpose of this project is to hold SQL skills to analyze user interactions and engagement with the Instagram application. As a data analyst working with the product team, the goal is to provide valuable insights to support decision-making within the business. The project involves answering specific questions posed by the management team related to marketing analysis and investor metrics.

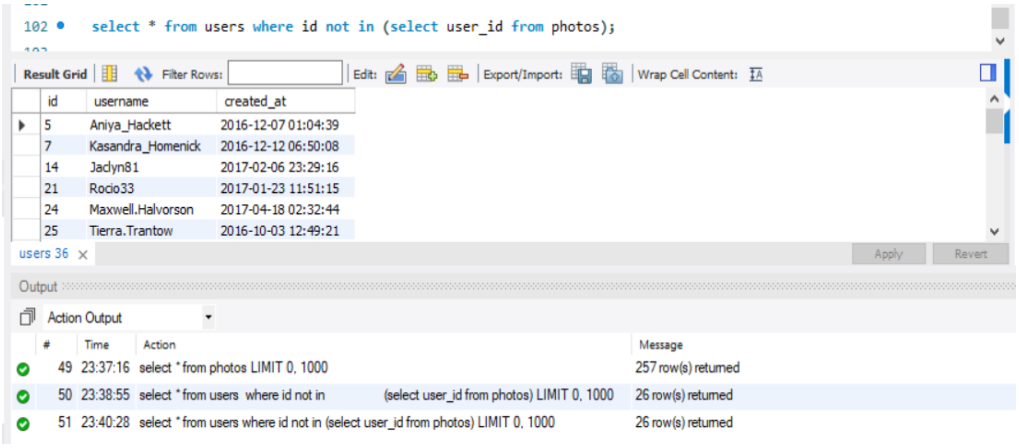
SQL Tasks :

A) Marketing Analysis:

1. **Loyal User Reward:** The marketing team wants to reward the most loyal users, i.e., those who have been using the platform for the longest time.
   Your Task: Identify the five oldest users on Instagram from the provided database.



2. **Inactive User Engagement:** The team wants to encourage inactive users to start posting by sending them promotional emails.
   Your Task: Identify users who have never posted a single photo on Instagram.



```
102 •  select * from users where id not in (select user_id from photos);
```

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

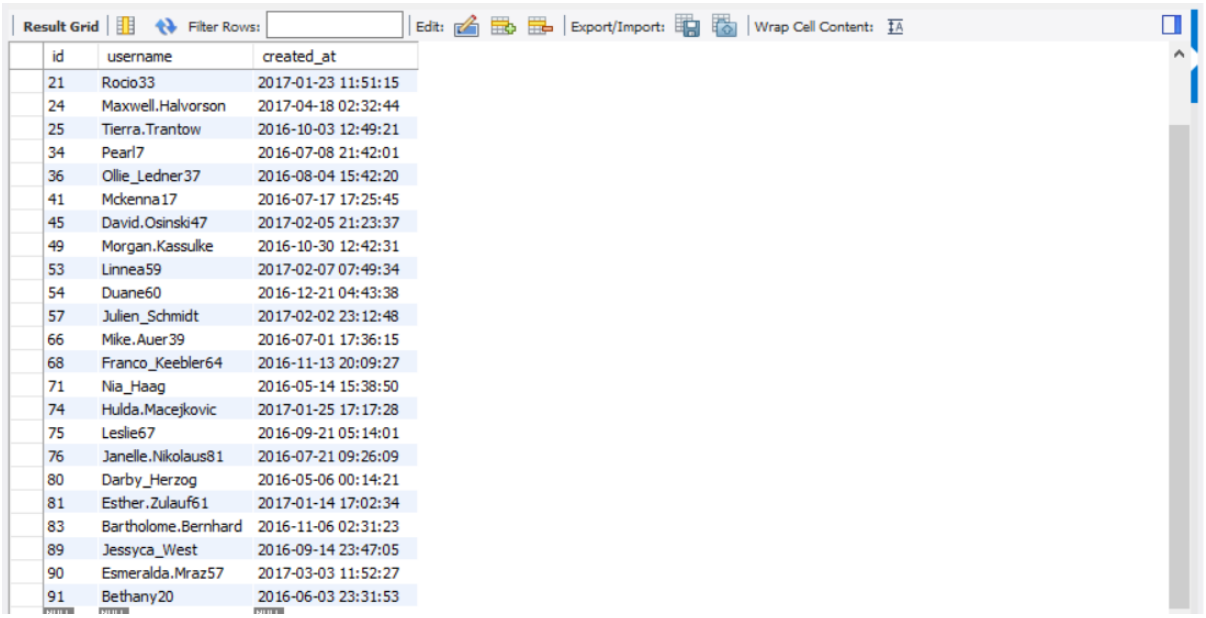
id	username	created_at
5	Aniya_Hackett	2016-12-07 01:04:39
7	Kasandra_Homenick	2016-12-12 06:50:08
14	Jadyn81	2017-02-06 23:29:16
21	Rocio33	2017-01-23 11:51:15
24	Maxwell.Halvorson	2017-04-18 02:32:44
25	Tierra.Trantow	2016-10-03 12:49:21

users 36 × | Apply | Revert

Output

Action Output

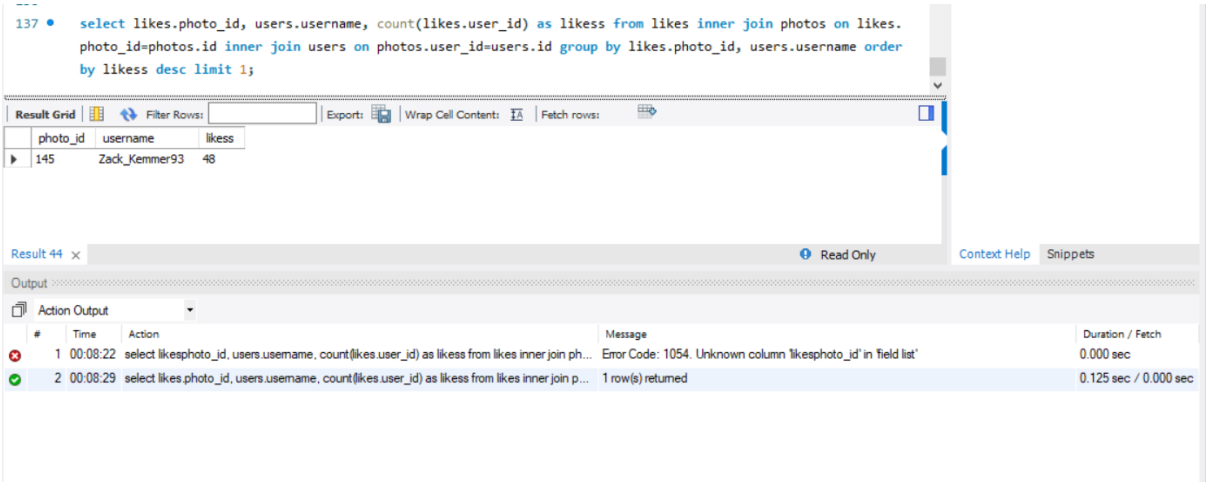
#	Time	Action	Message
49	23:37:16	select * from photos LIMIT 0, 1000	257 row(s) returned
50	23:38:55	select * from users where id not in (select user_id from photos) LIMIT 0, 1000	26 row(s) returned
51	23:40:28	select * from users where id not in (select user_id from photos) LIMIT 0, 1000	26 row(s) returned



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

id	username	created_at
21	Rocio33	2017-01-23 11:51:15
24	Maxwell.Halvorson	2017-04-18 02:32:44
25	Tierra.Trantow	2016-10-03 12:49:21
34	Pearl7	2016-07-08 21:42:01
36	Ollie_Ledner37	2016-08-04 15:42:20
41	Mckenna17	2016-07-17 17:25:45
45	David.Osinski47	2017-02-05 21:23:37
49	Morgan.Kassulke	2016-10-30 12:42:31
53	Linnea59	2017-02-07 07:49:34
54	Duane60	2016-12-21 04:43:38
57	Julien_Schmidt	2017-02-02 23:12:48
66	Mike.Auer39	2016-07-01 17:36:15
68	Franco_Keebler64	2016-11-13 20:09:27
71	Nia_Haag	2016-05-14 15:38:50
74	Hulda.Macejkovic	2017-01-25 17:17:28
75	Leslie67	2016-09-21 05:14:01
76	Janelle.Nikolaus81	2016-07-21 09:26:09
80	Darby_Herzog	2016-05-06 00:14:21
81	Esther.Zulauf61	2017-01-14 17:02:34
83	Bartholome.Bernhard	2016-11-06 02:31:23
89	Jessyca_West	2016-09-14 23:47:05
90	Esmeralda.Mraz57	2017-03-03 11:52:27
91	Bethany20	2016-06-03 23:31:53

3. **Contest Winner Declaration:** The team has organized a contest where the user with the most likes on a single photo wins.
   Your Task: Determine the winner of the contest and provide their details to the team.



```
137 •  select likes.photo_id, users.username, count(likes.user_id) as likess from likes inner join photos on likes.
       photo_id=photos.id inner join users on photos.user_id=users.id group by likes.photo_id, users.username order
       by likess desc limit 1;
```

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

photo_id	username	likess
145	Zack_Kemmer93	48

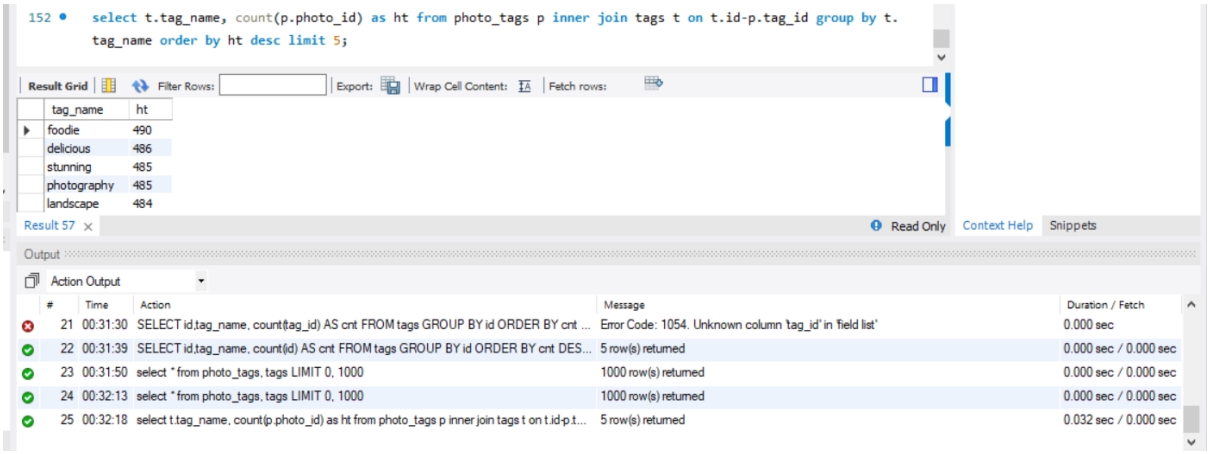
Result 44 | Read Only | Context Help | Snippets

Output

Action Output

#	Time	Action	Message	Duration / Fetch
1	00:08:22	select likesphoto_id, users.username, count(likes.user_id) as likess from likes inner join ph...	Error Code: 1054. Unknown column 'likesphoto_id' in 'field list'	0.000 sec
2	00:08:29	select likes.photo_id, users.username, count(likes.user_id) as likess from likes inner join p...	1 row(s) returned	0.125 sec / 0.000 sec

4. **Hashtag Research:** A partner brand wants to know the most popular hashtags to use in their posts to reach the most people.
   Your Task: Identify and suggest the top five most commonly used hashtags on the platform.



```
152 •  select t.tag_name, count(p.photo_id) as ht from photo_tags p inner join tags t on t.id-p.tag_id group by t.
       tag_name order by ht desc limit 5;
```

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

tag_name	ht
foodie	490
delicious	486
stunning	485
photography	485
landscape	484

Result 57 | Read Only | Context Help | Snippets

Output

Action Output

#	Time	Action	Message	Duration / Fetch
21	00:31:30	SELECT id,tag_name, count(tag_id) AS cnt FROM tags GROUP BY id ORDER BY cnt ...	Error Code: 1054. Unknown column 'tag_id' in 'field list'	0.000 sec
22	00:31:39	SELECT id,tag_name, count(id) AS cnt FROM tags GROUP BY id ORDER BY cnt DES...	5 row(s) returned	0.000 sec / 0.000 sec
23	00:31:50	select * from photo_tags, tags LIMIT 0, 1000	1000 row(s) returned	0.000 sec / 0.000 sec
24	00:32:13	select * from photo_tags, tags LIMIT 0, 1000	1000 row(s) returned	0.000 sec / 0.000 sec
25	00:32:18	select t.tag_name, count(p.photo_id) as ht from photo_tags p inner join tags t on t.id-p.t...	5 row(s) returned	0.032 sec / 0.000 sec

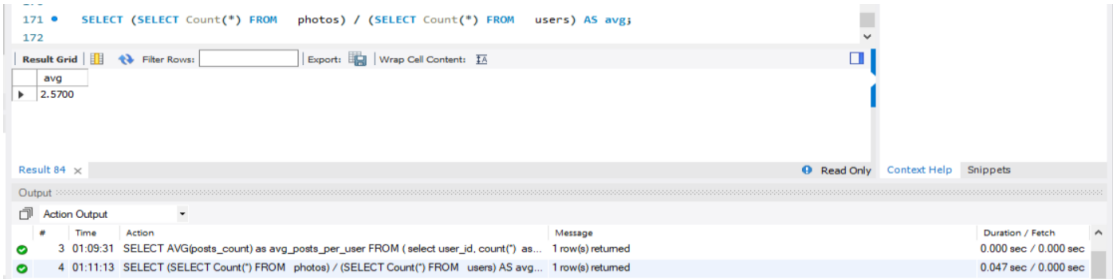
5. **Ad Campaign Launch:** The team wants to know the best day of the week to launch ads.
   Your Task: Determine the day of the week when most users register on Instagram. Provide insights on when to schedule an ad campaign.

B) Investor Metrics:

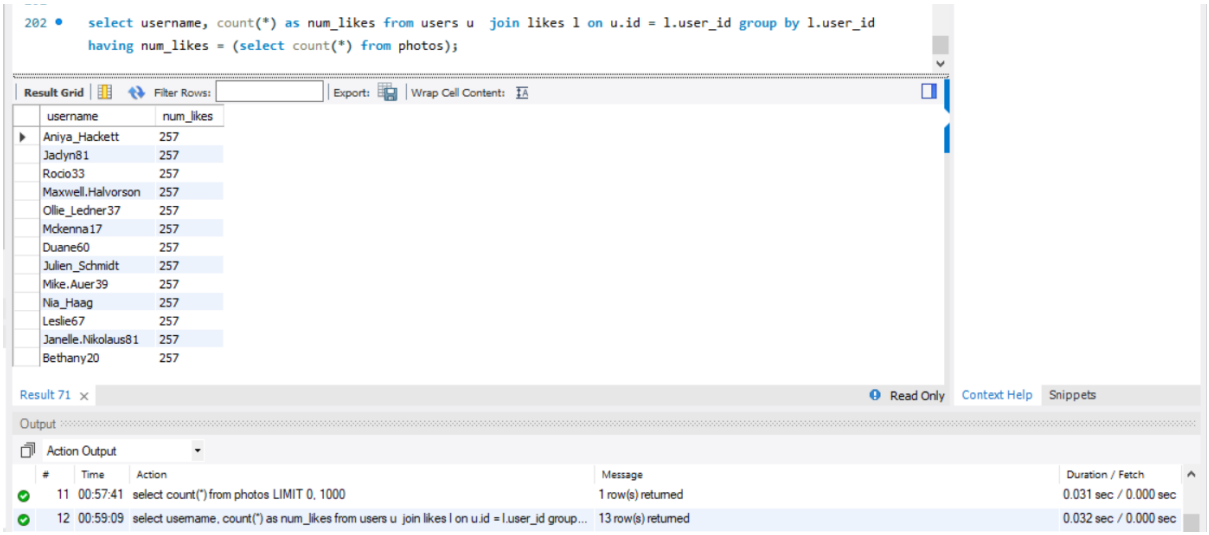
1. **User Engagement:** Investors want to know if users are still active and posting on Instagram or if they are making fewer posts.
   Your Task: Calculate the average number of posts per user on Instagram. Also, provide the total number of photos on Instagram divided by the total number of users.

Calculate the average number of posts per user on Instagram

the total number of photos on Instagram divided by the total number of users.



2. **Bots & Fake Accounts:** Investors want to know if the platform is crowded with fake and dummy accounts.
   Your Task: Identify users (potential bots) who have liked every single photo on the site, as this is not typically possible for a normal user.



```
202 •  select username, count(*) as num_likes from users u  join likes l on u.id = l.user_id group by l.user_id
       having num_likes = (select count(*) from 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

Result 71 × | Read Only | Context Help | Snippets

Output

Action Output

#	Time	Action	Message	Duration / Fetch
11	00:57:41	select count(*) from photos LIMIT 0, 1000	1 row(s) returned	0.031 sec / 0.000 sec
12	00:59:09	select username, count(*) as num_likes from users u join likes l on u.id = l.user_id group...	13 row(s) returned	0.032 sec / 0.000 sec

- **Project Description:-**

  project description outlines a scenario where you, as a data analyst working with the product team at Instagram, are tasked with analyzing user interactions and engagement using SQL and MySQL Workbench. The goal is to provide valuable insights to assist the product team in making informed decisions about the future direction of the Instagram app.

- **Approach:-**

  Database Creation:
  The Executed provided commands to create the necessary database. Therefore, Ensured proper structure and data representation.

  Marketing Analysis:
  - Identified the five oldest users using registration dates.
  - Identified inactive users by checking post counts.
  - Determined contest winners based on likes.
  - Analyzed post hashtags to suggest top five.
  - Determined the best day for ad campaign launches based on user registration.

Investor Metrics:
by Calculated average posts per user and total photos divided by total users.
Identified potential bots by finding users liking every photo.

Report Preparation:
Documented SQL queries along with outputs.
Created a report in PDF format.

- **Tech-Stack Used:- Tech-Stack Used:**

  MySQL Workbench (Version 8.0 CE):

  Chosen for its user-friendly interface and powerful SQL query capabilities.

  Provides a seamless environment for database creation, management, and analysis.

- **Insights:**
  User Engagement:
- Average posts per user indicated healthy user activity.
- Understanding user engagement is crucial for decision-making in feature development and marketing strategies.

- **Bots & Fake Accounts:**
- Identified potential bots to maintain a genuine user base.
- Addressing fake accounts contributes to a more reliable platform.

- **Marketing Insights:**
- Recognized the loyalty of the oldest users for possible loyalty programs.
- Identified inactive users for targeted re-engagement strategies.
- Declared contest winners and suggested popular hashtags for marketing campaigns.

- **Ad Campaign Optimization:**
- Determined the best day for ad campaign launches based on user registration patterns.

- **Result:-**

  Achievements:

  - Successfully completed all tasks, providing valuable insights for decision-making.

- how as a business or data analyst we work on real-time data to take any data-driven decision.
- Achieved a comprehensive understanding of user behavior.

Impact:

- The analysis contributes to informed decision-making for product development, marketing, and user engagement.
- Insights derived from the project can potentially influence the future development of Instagram.

Drive Link

By following this approach, the project showcases the application of SQL skills to extract meaningful insights from Instagram user data, contributing to the strategic direction of one of the world's most popular social media platforms.