Instagram User Analytics

SQL Fundamentals

Instagram User Analytics

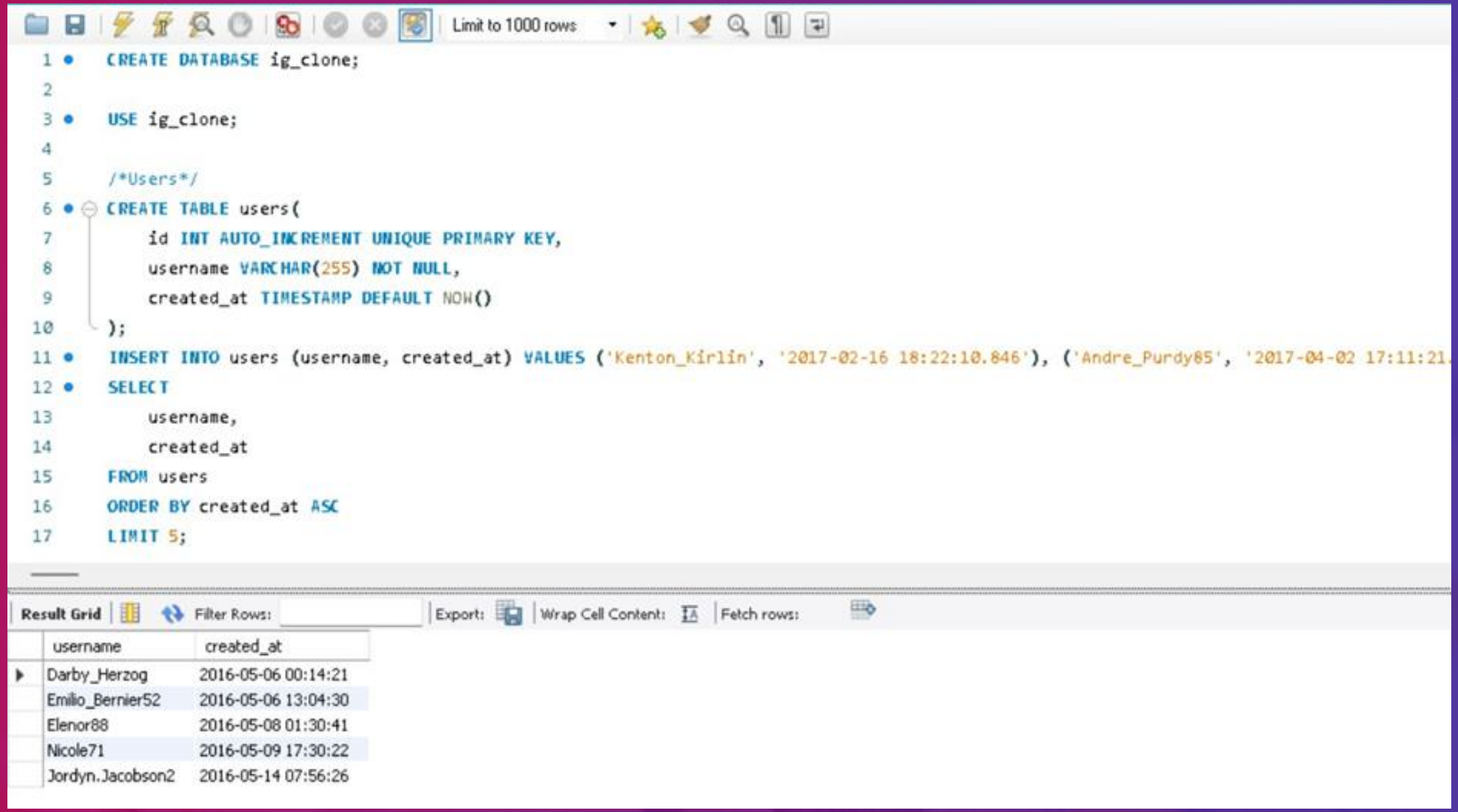
Using SQL and MySQL Workbench

In this project, I am utilizing SQL and MySQL Workbench to analyze Instagram user data and address questions from the management team. These insights will assist the product manager and the rest of the team in making informed decisions about the future direction of the Instagram app.

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.

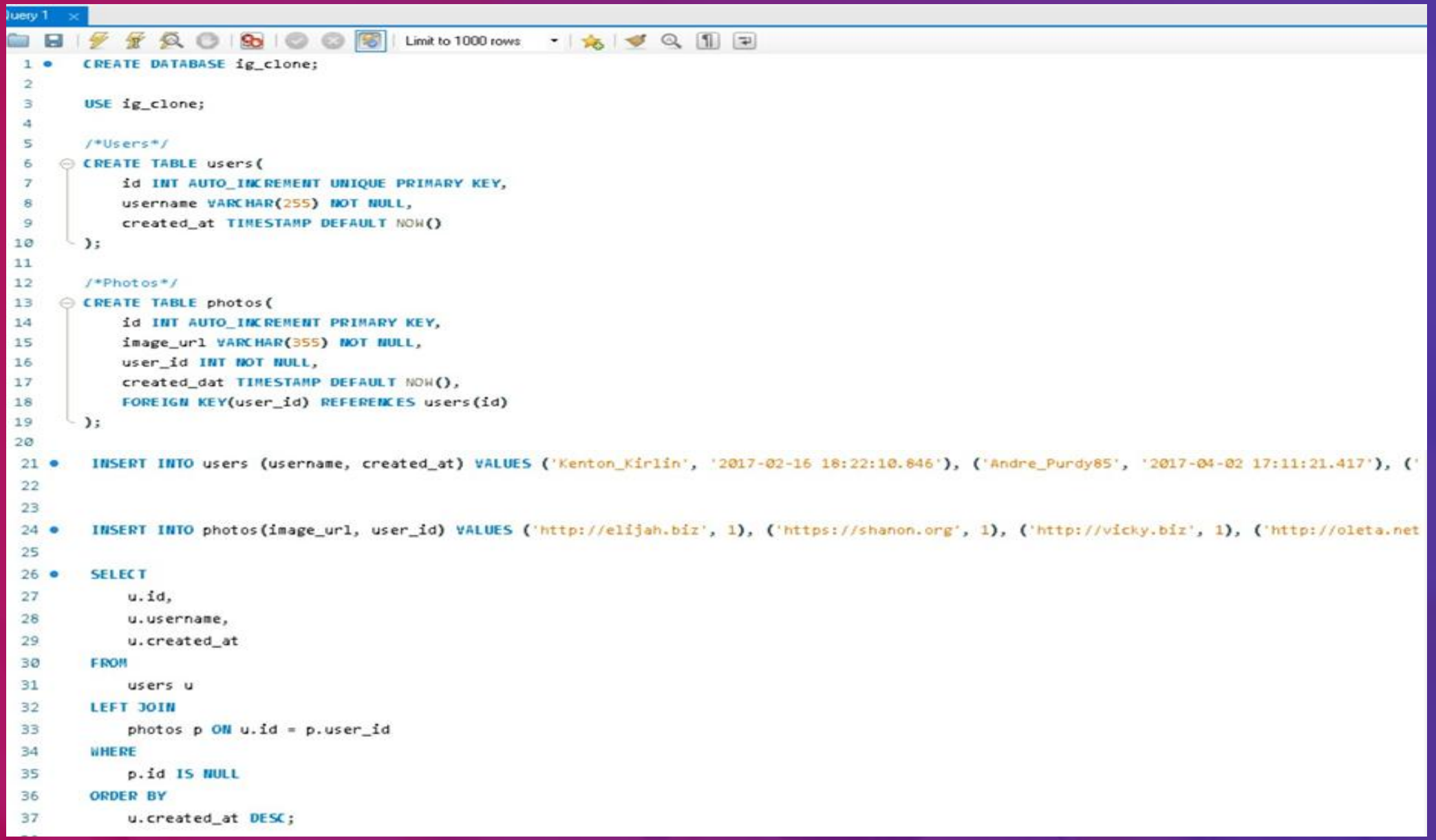


The five oldest users on Instagram

Marketing Analysis:

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.



Users who have never posted a single photo on Instagram - Query

Result Grid | Filter Rows: | Exp

id	username	created_at
5	Aniya_Hackett	2016-12-07 01:04:39
7	Kasandra_Homenick	2016-12-12 06:50:08
14	Jaclyn81	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
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

Result Grid | Filter Rows:

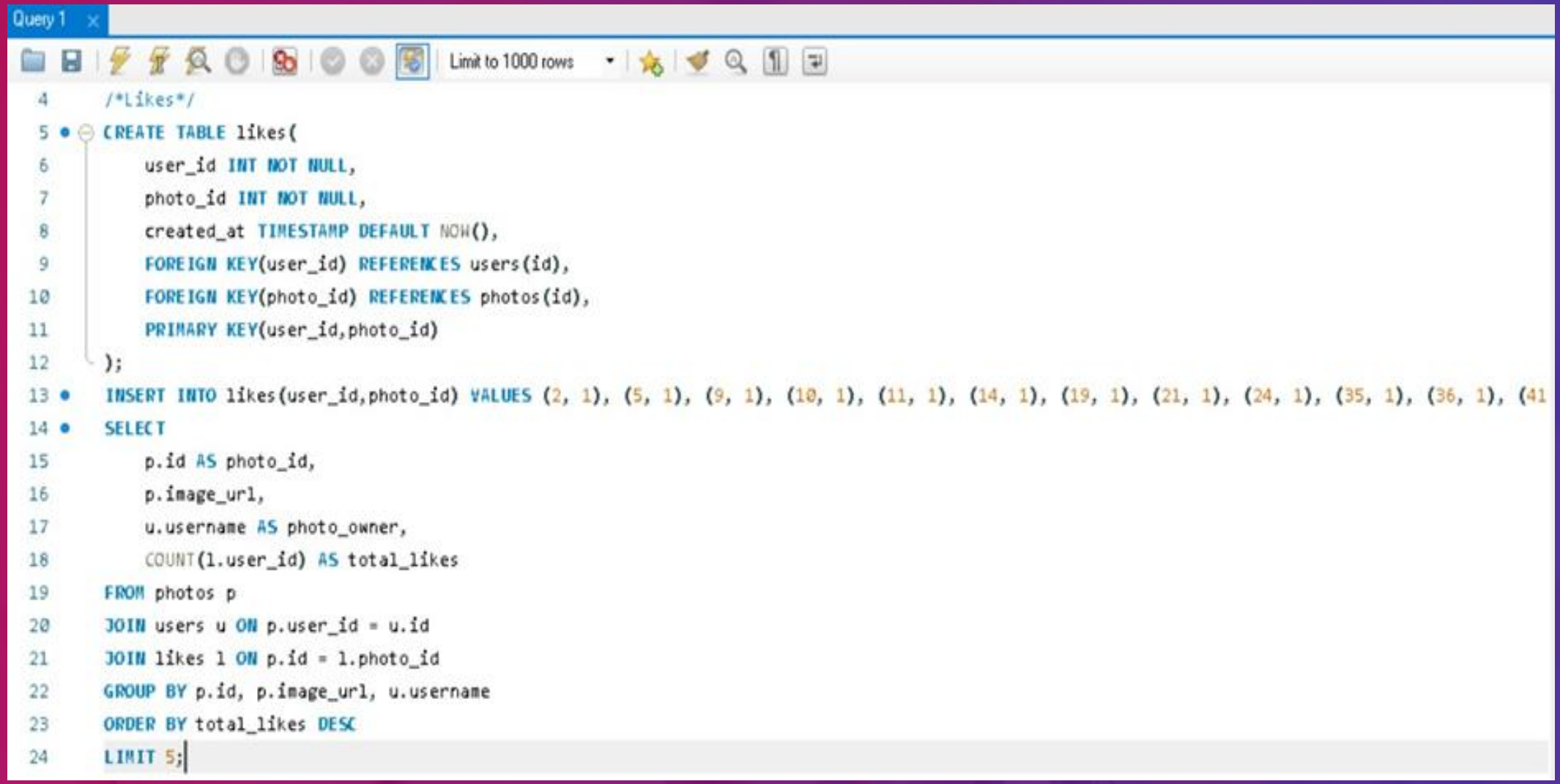
id	username	created_at
57	Julien_Schmidt	2017-02-02 23:12:48
66	Mike.Auer39	2016-07-01 17:36:15
68	Franco_Keebler64	2016-11-13 20:09:27
71	Nia_Haag	2016-05-14 15:38:50
74	Hulda.Macejkovic	2017-01-25 17:17:28
75	Leslie67	2016-09-21 05:14:01
76	Janelle.Nikolaus81	2016-07-21 09:26:09
80	Darby_Herzog	2016-05-06 00:14:21
81	Esther.Zulauf61	2017-01-14 17:02:34
83	Bartholome.Bernh...	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

Users who have never posted a single photo on Instagram - Result

Marketing Analysis:

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.



The winner of the contest and their details - Query

Result Grid | Filter Rows: | Export:

photo_id	image_url	photo_owner	total_likes
145	https://jarret.name	Zack_Kemmer93	48
127	https://celestine.name	Malinda_Streich	43
182	https://dorcas.biz	Adelle96	43
123	http://shannon.org	Seth46	42
61	https://dejon.name	Delpha.Kihn	41

The winner of the contest and their details - Result

Marketing Analysis:

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.

```sql
CREATE DATABASE ig_clone;

USE ig_clone;

/*Tags*/
CREATE TABLE tags(
    id INTEGER AUTO_INCREMENT PRIMARY KEY,
    tag_name VARCHAR(255) UNIQUE NOT NULL,
    created_at TIMESTAMP DEFAULT NOW()
);

/*junction table: Photos - Tags*/
CREATE TABLE photo_tags(
    photo_id INT NOT NULL,
    tag_id INT NOT NULL,
    FOREIGN KEY(photo_id) REFERENCES photos(id),
    FOREIGN KEY(tag_id) REFERENCES tags(id),
    PRIMARY KEY(photo_id,tag_id)
);
INSERT INTO tags(tag_name) VALUES ('sunset'), ('photography'), ('sunrise'), ('landscape'), ('food'), ('foodie'), ('delicious'), ('beauty'), ('



INSERT INTO photo_tags(photo_id, tag_id) VALUES (1, 18), (1, 17), (1, 21), (1, 13), (1, 19), (2, 4), (2, 3), (2, 20), (2, 2), (3, 8), (4, 12),

SELECT
   t.tag_name AS "Tag Name",
   COUNT(*) AS "Usage Count"
FROM photo_tags pt
JOIN tags t ON pt.tag_id = t.id
GROUP BY t.tag_name
ORDER BY COUNT(*) DESC
LIMIT 5;
```

Result Grid | Filter Rows:

Tag Name	Usage Count
smile	59
party	39
fun	38
concert	24
beach	42

The top five most commonly used hashtags

Marketing Analysis:

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.

```sql
CREATE DATABASE ig_clone;

USE ig_clone;

/*Users*/
CREATE TABLE users(
    id INT AUTO_INCREMENT UNIQUE PRIMARY KEY,
    username VARCHAR(255) NOT NULL,
    created_at TIMESTAMP DEFAULT NOW()
);

INSERT INTO users (username, created_at) VALUES ('Kenton_Kirlin', '2017-02-16 18:22:10.846'), ('Andre_Purdy85', '2017-04-02 17:11:21.417'), ('

SELECT
    DAYNAME(created_at) AS day_of_week,
    COUNT(*) AS registrations
FROM users
GROUP BY DAYNAME(created_at)
ORDER BY registrations DESC;
```

Result Grid | Filter Rows:

day_of_week	registrations
Thursday	16
Sunday	16
Friday	15
Tuesday	14
Monday	14
Wednesday	13
Saturday	12

The days of the week when most users register on Instagram

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.

Query 1

Limit to 1000 rows

```sql
CREATE DATABASE ig_clone;

USE ig_clone;

/*Photos*/
CREATE TABLE photos(
    id INT AUTO_INCREMENT PRIMARY KEY,
    image_url VARCHAR(355) NOT NULL,
    user_id INT NOT NULL,
    created_dat TIMESTAMP DEFAULT NOW(),
    FOREIGN KEY(user_id) REFERENCES users(id)
);

INSERT INTO photos(image_url, user_id) VALUES ('http://elijah.biz', 1), ('https://shanon.org', 1), ('http://vicky.biz', 1), ('http://oleta.net

SELECT
    ROUND(COUNT(*) / COUNT(DISTINCT user_id), 2) as avg_posts_per_active_user,
    COUNT(*) as total_photos,
    COUNT(DISTINCT user_id) as users_with_posts
FROM photos;
```

Result Grid | Filter Rows: | Export:

avg_posts_per_active_user	total_photos	users_with_posts
3.47	257	74

The average number of posts per user on Instagram

Investor Metrics:

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.

Query 1

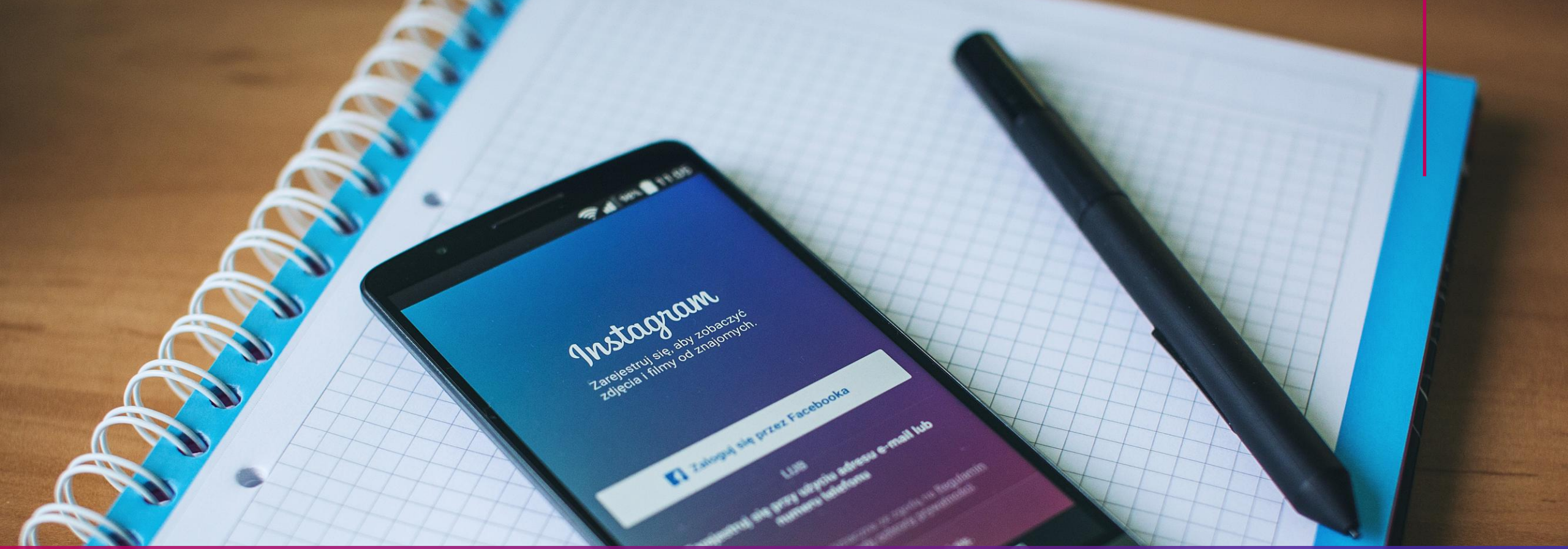
Limit to 1000 rows

```sql
CREATE DATABASE ig_clone;
USE ig_clone;

/*Photos*/
CREATE TABLE photos(
    id INT AUTO_INCREMENT PRIMARY KEY,
    image_url VARCHAR(355) NOT NULL,
    user_id INT NOT NULL,
    created_dat TIMESTAMP DEFAULT NOW(),
    FOREIGN KEY(user_id) REFERENCES users(id)
);

/*Likes*/
CREATE TABLE likes(
    user_id INT NOT NULL,
    photo_id INT NOT NULL,
    created_at TIMESTAMP DEFAULT NOW(),
    FOREIGN KEY(user_id) REFERENCES users(id),
    FOREIGN KEY(photo_id) REFERENCES photos(id),
    PRIMARY KEY(user_id,photo_id)
);


INSERT INTO photos(image_url, user_id) VALUES ('http://elijah.biz', 1), ('https://shanon.org', 1), ('http://vicky.biz', 1), ('http://oleta.net

INSERT INTO likes(user_id,photo_id) VALUES (2, 1), (5, 1), (9, 1), (10, 1), (11, 1), (14, 1), (19, 1), (21, 1), (24, 1), (35, 1), (36, 1), (41

SELECT
    l.user_id
FROM
    likes l
GROUP BY l.user_id
HAVING COUNT(DISTINCT l.photo_id) = (SELECT
        COUNT(*)
    FROM
        photos);
```

Result Grid

user_id
5
14
21
24
36
41
54
57
66
71
75
76
91

Users (potential bots) who have liked every single photo on the site


Instagram
Zarejestruj się, aby zobaczyć zdjęcia i filmy od znajomych.
Zaloguj się przez Facebooka
LUB

THANK YOU!