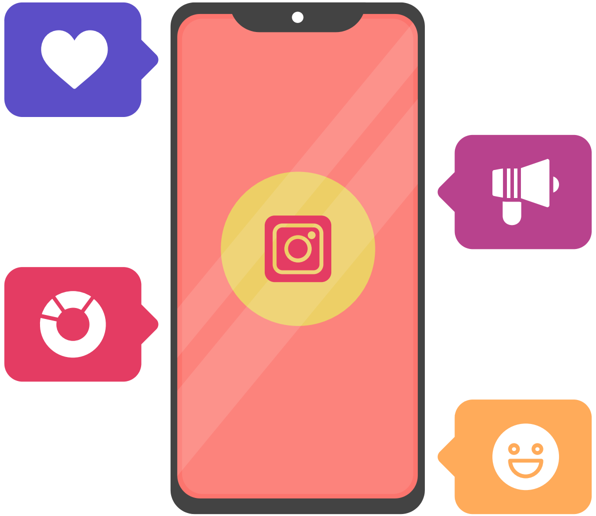


Instagram User Analytics

Trainity Project Report

Rohit Kumar
rohitk.ug20.ce@nitp.ac.in

Description

This report analyses how consumers engage and interact with our digital platform (software or mobile application) to gain marketing and investor metrics business insights. Teams from throughout the company can utilize these insights to start a new marketing campaign, choose which features to include in an app, gauge the app's success by looking at user engagement, and generally improve the user experience while supporting the expansion of the company.

Approach

This report will be divided up primarily into two sections.

1) Marketing

I. **Rewarding Most Loyal Users** - To find most loyal users – for rewarding purpose (Top 5 oldest users of the Instagram)

II. **Remind Inactive Users to Start Posting** – Sending mail to those users who have never posted single photo yet

III. **Declaring Contest Winner** -To find details of Winner(user) of contest who gets most likes on single photo

IV. **Hashtag Researching** - To find top 5 Hashtag- for better reach to the most people

V. **Launch AD Campaign** - To find the day on which users register more Instagram

Approach

2) Investor Metrics

I. User Engagement - Are users still active on Instagram or not?

- To find this I will check

a) how many times does average user posts on Instagram?

b) Total number of photos on Instagram/total number of users.

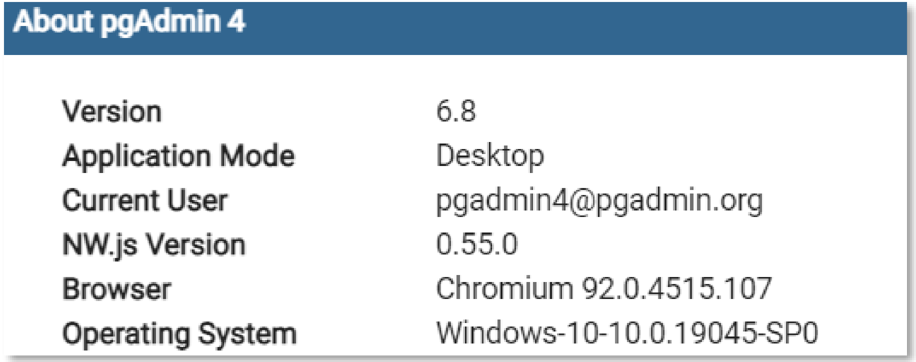
II. Bots & Fake Accounts - Are there any fake/dummy accounts are active on Instagram?

- To find this I will check

a) Data on users (bots) who have liked every single photo on the site (since any normal user would not be able to do this).

Tech-Stack Used

- For this Project I used **PostgreSQL PG Admin 4 version 6.8**



About pgAdmin 4

Version	6.8
Application Mode	Desktop
Current User	pgadmin4@pgadmin.org
NW.js Version	0.55.0
Browser	Chromium 92.0.4515.107
Operating System	Windows-10-10.0.19045-SP0

- This project is based on the provided "IG_clone" dataset which I modified due to syntax error. [IG_Clone DataSet]

ANALYSIS & RESULT

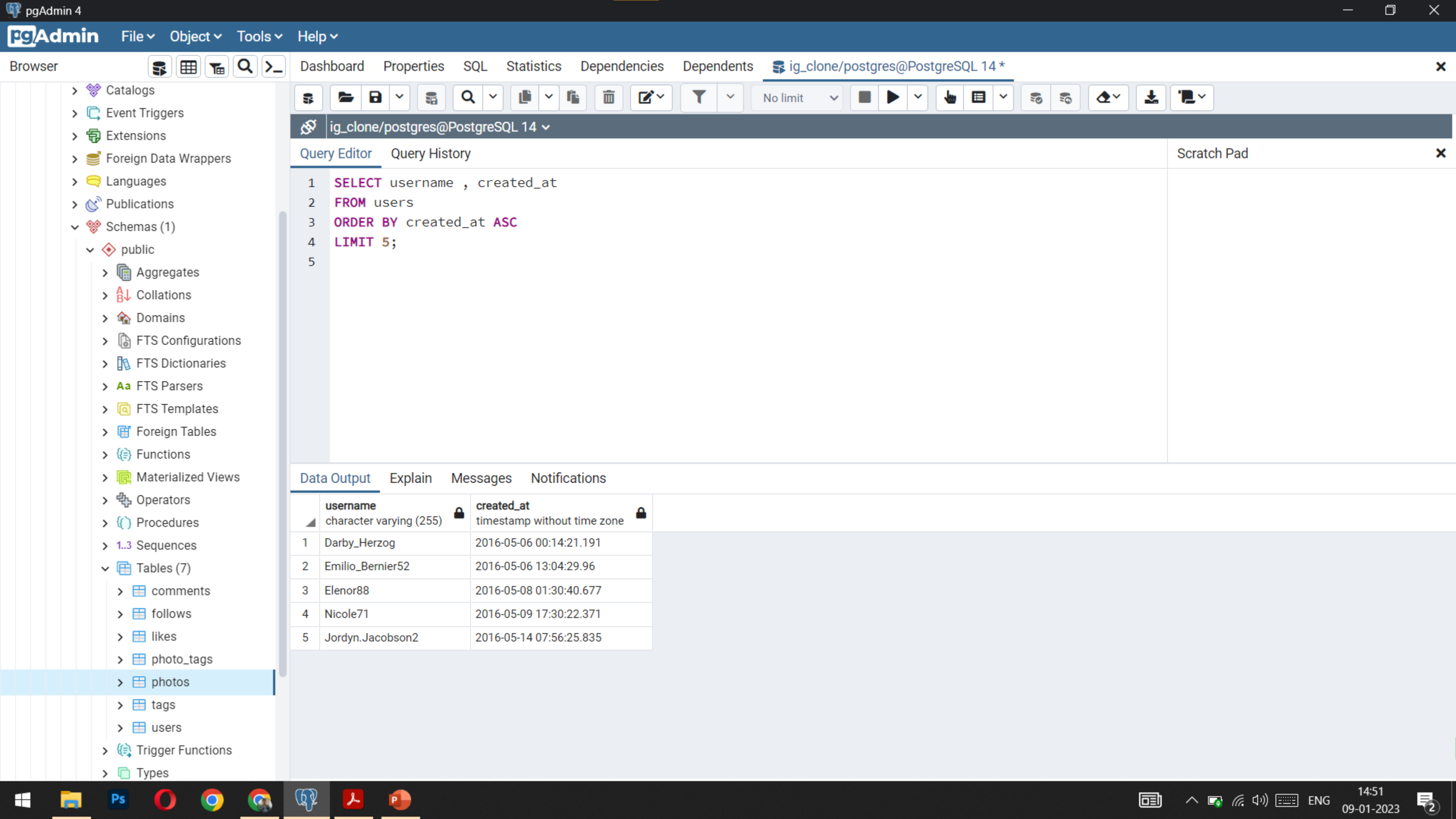
1 (I) Rewarding Most Loyal Users

- Top 5 oldest users of the Instagram are here below

1. Darby_Herzog
2. Emilio_Bernier52
3. Elenor88
4. Nicole71
5. Jordyn.Jacobson2

Query:-

```
SELECT username , created_at
FROM users
ORDER BY created_at ASC
LIMIT 5;
```



```sql
SELECT username , created_at
FROM users
ORDER BY created_at ASC
LIMIT 5;
```

	username character varying (255)	created_at timestamp without time zone
1	Darby_Herzog	2016-05-06 00:14:21.191
2	Emilio_Bernier52	2016-05-06 13:04:29.96
3	Elenor88	2016-05-08 01:30:40.677
4	Nicole71	2016-05-09 17:30:22.371
5	Jordyn.Jacobson2	2016-05-14 07:56:25.835

1 (II) Remind Inactive Users to Start Posting

- List of inactive users or users who never post single post yet are here below

ID	USERNAME
5	"Aniya_Hackett"
7	"Kasandra_Homenick"
14	"Jaclyn81"
21	"Rocio33"
24	"Maxwell.Halvorson"
25	"Tierra.Trantow"
34	"Pearl7"
36	"Ollie_Ledner37"
41	"Mckenna17"

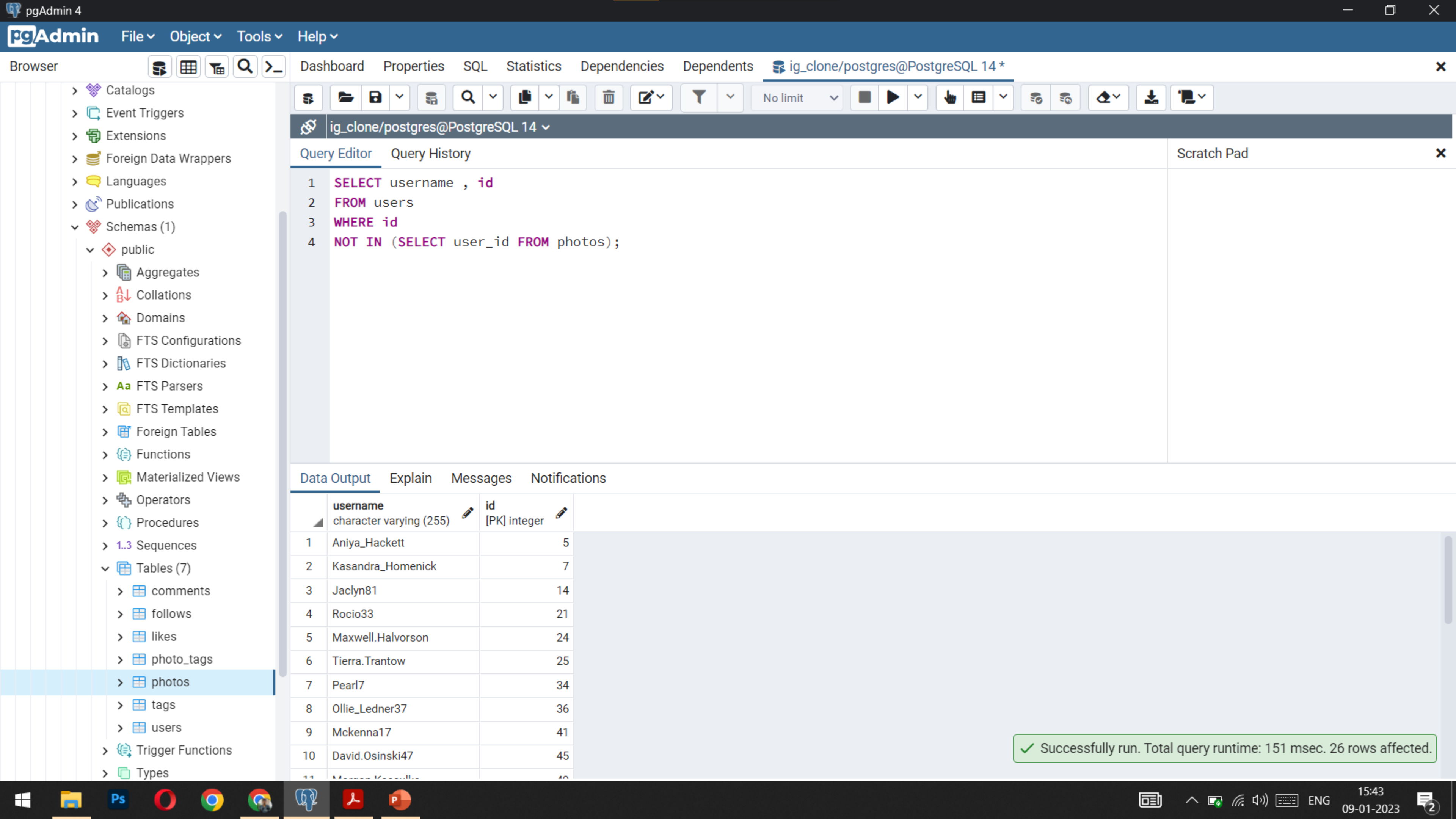
ID	USERNAME
45	"David.Osinski47"
49	"Morgan.Kassulke"
53	"Linnea59"
54	"Duane60"
57	"Julien_Schmidt"
66	"Mike.Auer39"
68	"Franco_Keebler64"
71	"Nia_Haag"
74	"Hulda.Macejkovic"

ID	USERNAME
75	"Leslie67"
76	"Janelle.Nikolaus81"
80	"Darby_Herzog"
81	"Esther.Zulauf61"
83	"Bartholome.Bernhard"
89	"Jessyca_West"
90	"Esmeralda.Mraz57"
91	"Bethany20"

1 (II) Remind Inactive Users to Start Posting

Query:-

```
SELECT username , id
FROM users
WHERE id
NOT IN (SELECT user_id FROM photos);
```



```sql
SELECT username , id
FROM users
WHERE id
NOT IN (SELECT user_id FROM photos);
```

	username character varying (255)	id [PK] integer
1	Aniya_Hackett	5
2	Kasandra_Homenick	7
3	Jaclyn81	14
4	Rocio33	21
5	Maxwell.Halvorson	24
6	Tierra.Trantow	25
7	Pearl7	34
8	Ollie_Ledner37	36
9	Mckenna17	41
10	David.Osinski47	45

✓ Successfully run. Total query runtime: 151 msec. 26 rows affected.

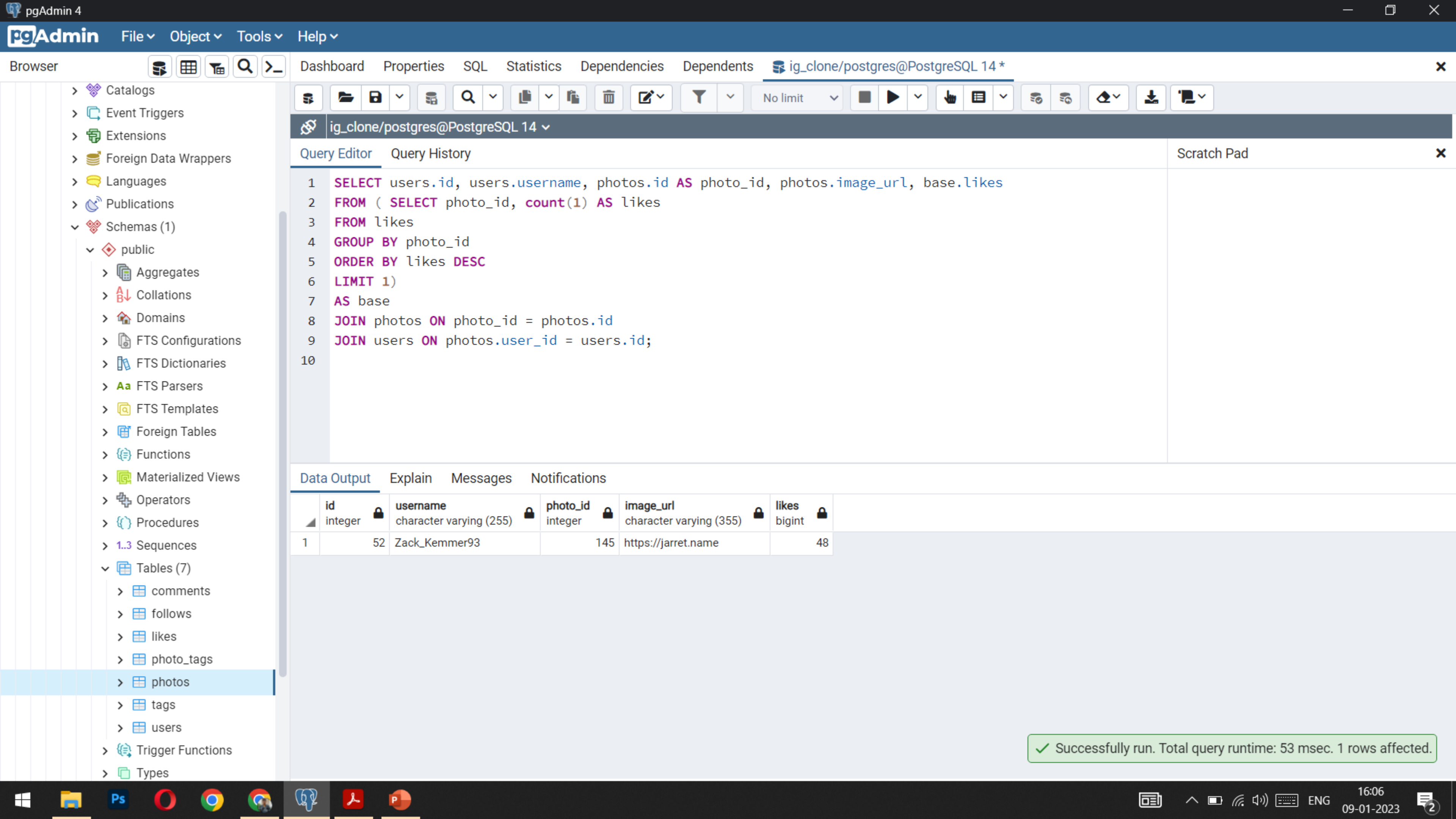
1 (III) Declaring Contest Winner

- Details of winner

ID	USERNAME	PHOTO_ID	IMAGE_URL	LIKES
52	"Zack_Kemmer93"	145	https://jarret.name	48

Query:-

```
SELECT users.id, users.username, photos.id AS photo_id, photos.image_url, base.likes
FROM ( SELECT photo_id, count(1) AS likes
FROM likes
GROUP BY photo_id
ORDER BY likes DESC
LIMIT 1)
AS base
JOIN photos ON photo_id = photos.id
JOIN users ON photos.user_id = users.id;
```



```sql
SELECT users.id, users.username, photos.id AS photo_id, photos.image_url, base.likes
FROM ( SELECT photo_id, count(1) AS likes
FROM likes
GROUP BY photo_id
ORDER BY likes DESC
LIMIT 1)
AS base
JOIN photos ON photo_id = photos.id
JOIN users ON photos.user_id = users.id;
```

	id integer	username character varying (255)	photo_id integer	image_url character varying (355)	likes bigint
1	52	Zack_Kemmer93	145	https://jarret.name	48

✓ Successfully run. Total query runtime: 53 msec. 1 rows affected.

1 (IV) Hashtag Researching

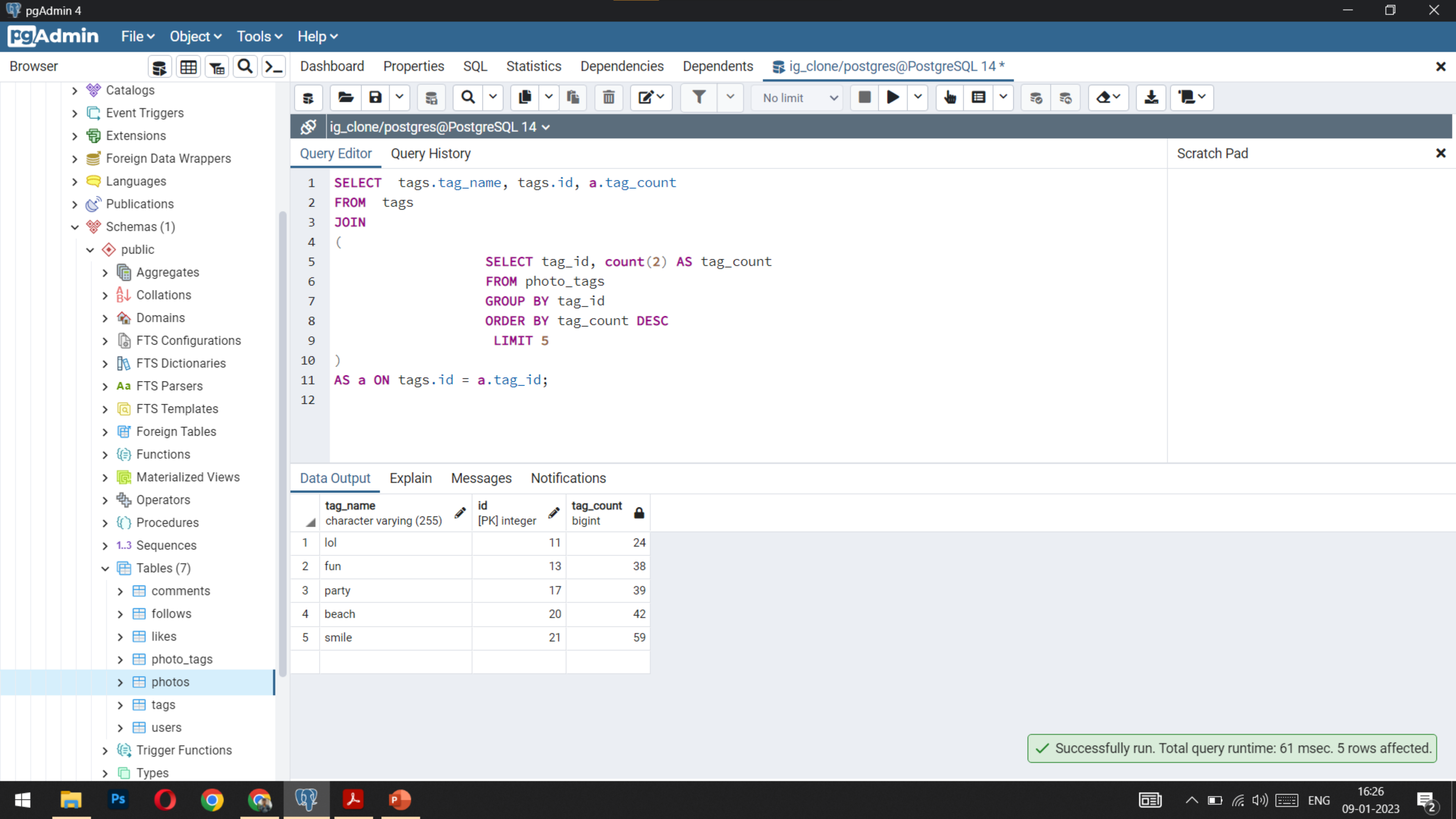
- Top 5 most used Hashtags

ID	TAG_NAME	TAG_COUNT
11	Lol	24
13	Fun	38
17	Party	39
20	Beach	42
21	Smile	59

1 (IV) Hashtag Researching

Query:-

```
SELECT  tags.tag_name, tags.id, a.tag_count
FROM  tags
JOIN
(
            SELECT tag_id, count(2) AS tag_count
            FROM photo_tags
            GROUP BY tag_id
            ORDER BY tag_count DESC
            LIMIT 5
)
AS a ON tags.id = a.tag_id;
```



```sql
SELECT  tags.tag_name, tags.id, a.tag_count
FROM  tags
JOIN
(
                   SELECT tag_id, count(2) AS tag_count
                   FROM photo_tags
                   GROUP BY tag_id
                   ORDER BY tag_count DESC
                    LIMIT 5
)
AS a ON tags.id = a.tag_id;
```

	tag_name character varying (255)	id [PK] integer	tag_count bigint
1	lol	11	24
2	fun	13	38
3	party	17	39
4	beach	20	42
5	smile	21	59

✓ Successfully run. Total query runtime: 61 msec. 5 rows affected.

1 (V) Launch Ad Campaign

- Day of the week do most users register on

Weekday No.	Weekday	User registered
4	Thursday	16
0	Sunday	16
5	Friday	15
1	Monday	14
2	Tuesday	14
3	Wednesday	13
6	Saturday	12

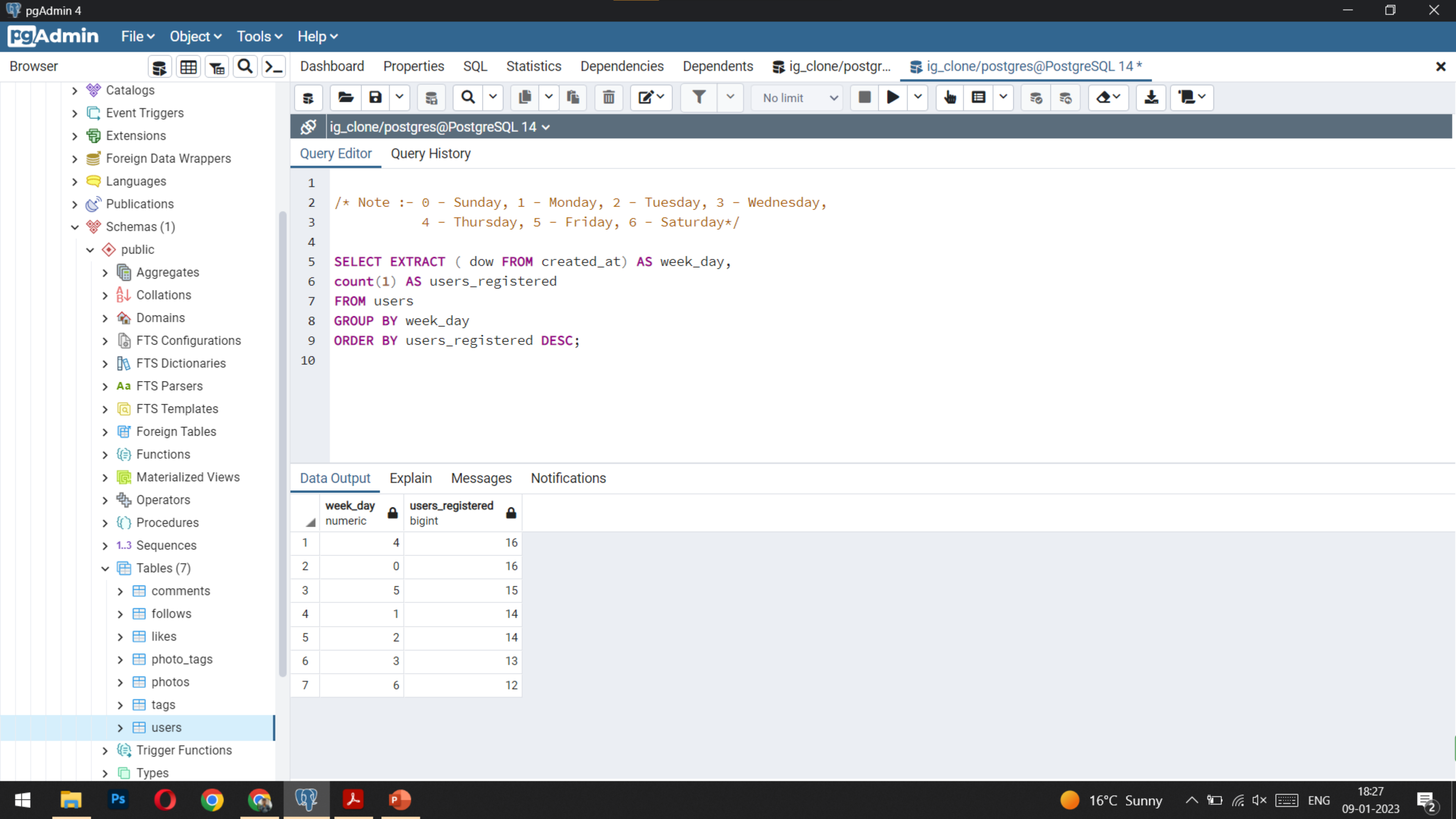
- On **Thursday** & **Sunday** user **registered** on Instagram is **maximum.**

1 (V) Launch Ad Campaign

Query:-

```
/* Note :- 0 - Sunday, 1 - Monday, 2 - Tuesday, 3 - Wednesday,
        4 - Thursday, 5 - Friday, 6 - Saturday*/

SELECT EXTRACT ( dow FROM created_at) AS week_day,
count(1) AS users_registered
FROM users
GROUP BY week_day
ORDER BY users_registered DESC;
```



```

/* Note :- 0 - Sunday, 1 - Monday, 2 - Tuesday, 3 - Wednesday,
           4 - Thursday, 5 - Friday, 6 - Saturday*/

SELECT EXTRACT ( dow FROM created_at) AS week_day,
count(1) AS users_registered
FROM users
GROUP BY week_day
ORDER BY users_registered DESC;

```

	week_day numeric	users_registered bigint
1	4	16
2	0	16
3	5	15
4	1	14
5	2	14
6	3	13
7	6	12

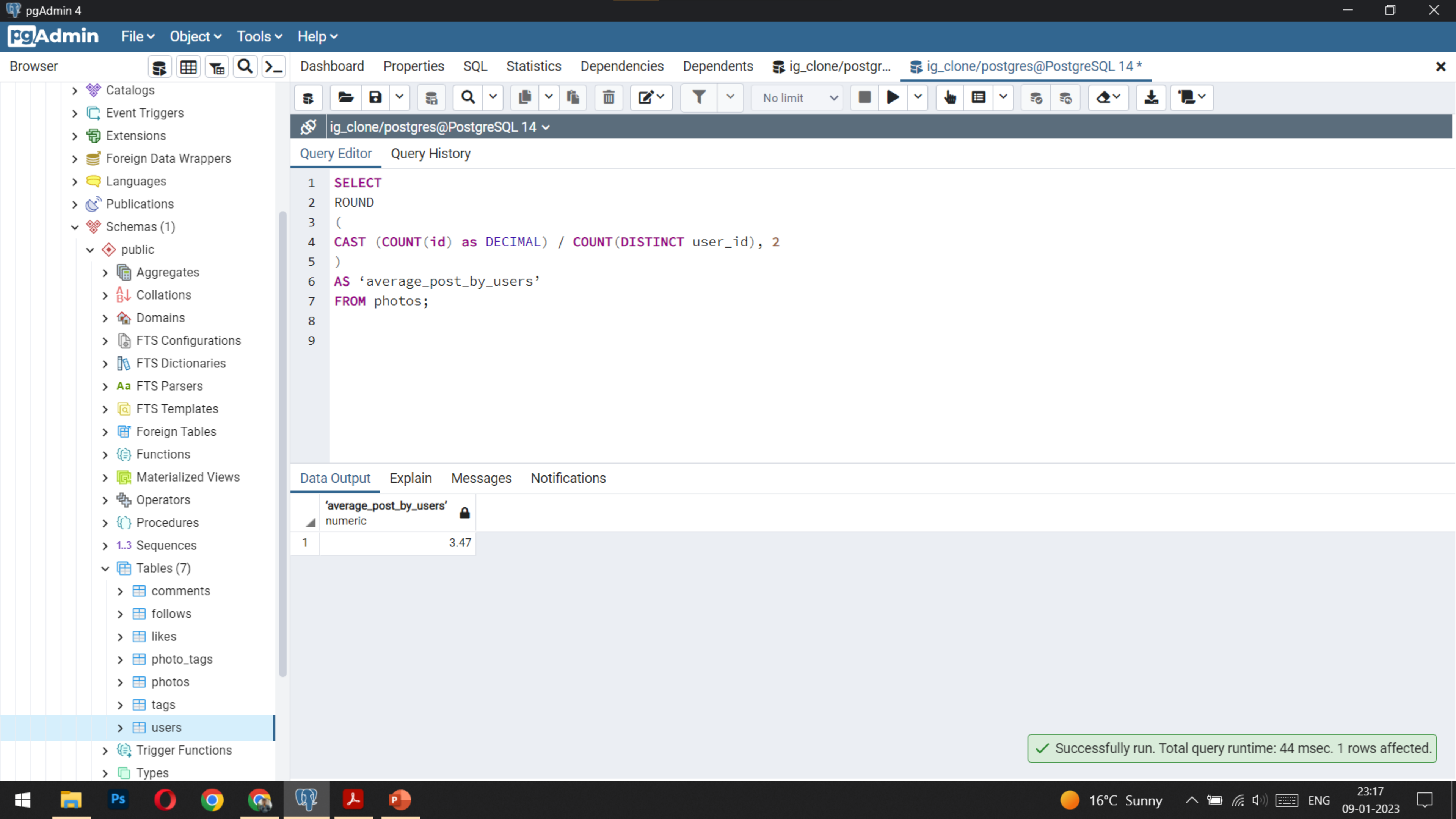
2 (I) User Engagement

a) How many times does average user posts on Instagram?

Ans: On average 3-4 user posts on Instagram.

Query:-

```
SELECT
ROUND
(
CAST (COUNT(id) as DECIMAL) / COUNT(DISTINCT user_id), 2
)
AS ‘average_post_by_users’
FROM photos;
```



```sql
SELECT
ROUND
(
CAST (COUNT(id) as DECIMAL) / COUNT(DISTINCT user_id), 2
)
AS ‘average_post_by_users’
FROM photos;
```

	‘average_post_by_users’ numeric
1	3.47

✓ Successfully run. Total query runtime: 44 msec. 1 rows affected.

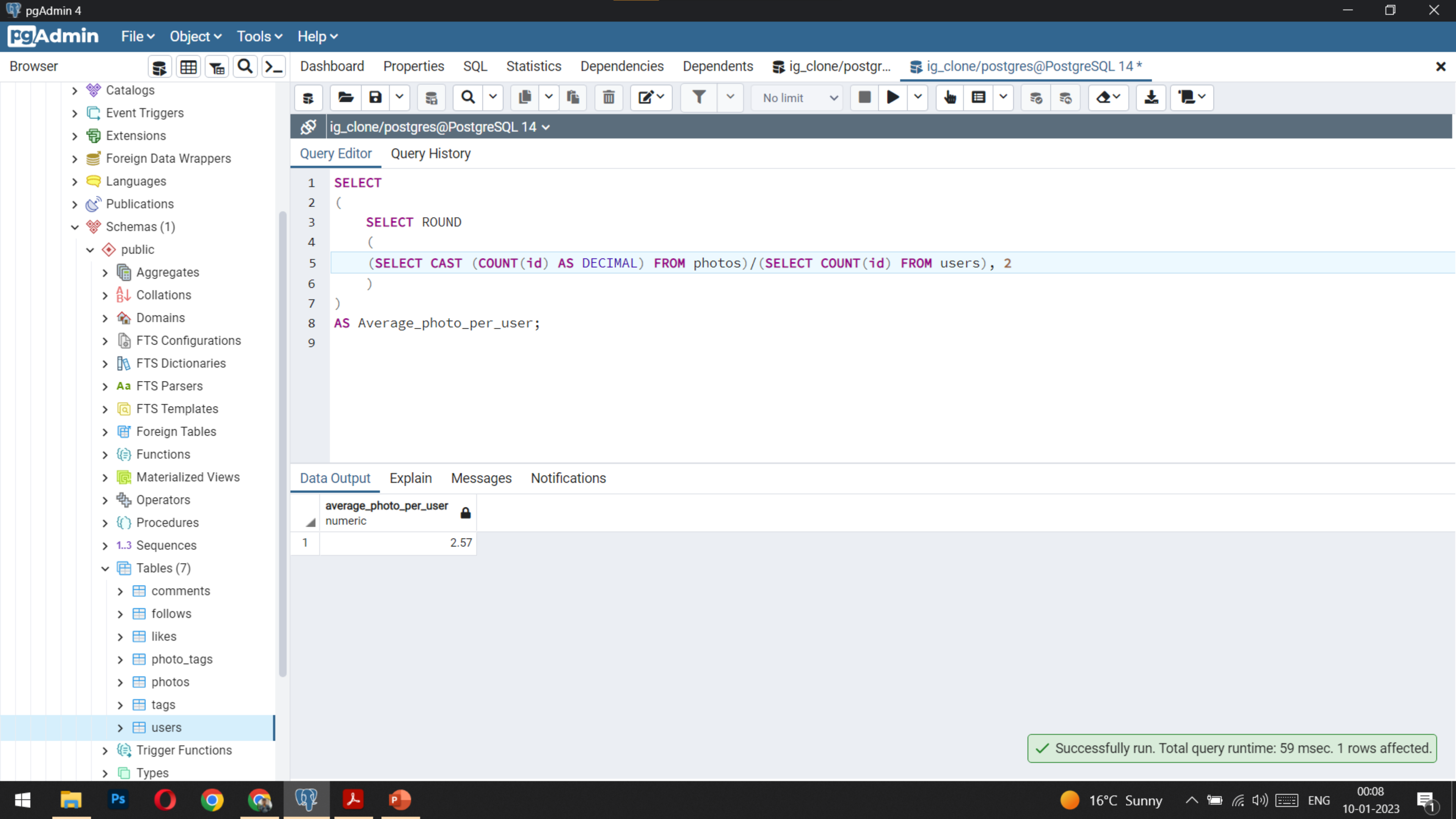
2 (I) User Engagement

b) Total number of photos on Instagram/total number of users.

Ans: 257/100 = 2.57

Query:-

```
SELECT
(
 SELECT ROUND
(  (SELECT CAST ( COUNT(id) AS DECIMAL) FROM photos) / (SELECT COUNT(id)
FROM users), 2 )
)
AS Average_photo_per_user;
```



```sql
SELECT
(
    SELECT ROUND
    (
    (SELECT CAST (COUNT(id) AS DECIMAL) FROM photos)/(SELECT COUNT(id) FROM users), 2
    )
)
AS Average_photo_per_user;
```

	average_photo_per_user numeric
1	2.57

✓ Successfully run. Total query runtime: 59 msec. 1 rows affected.

2 (II) Bots & Fake Accounts

- Bots/ Fake Accounts are as follows :-

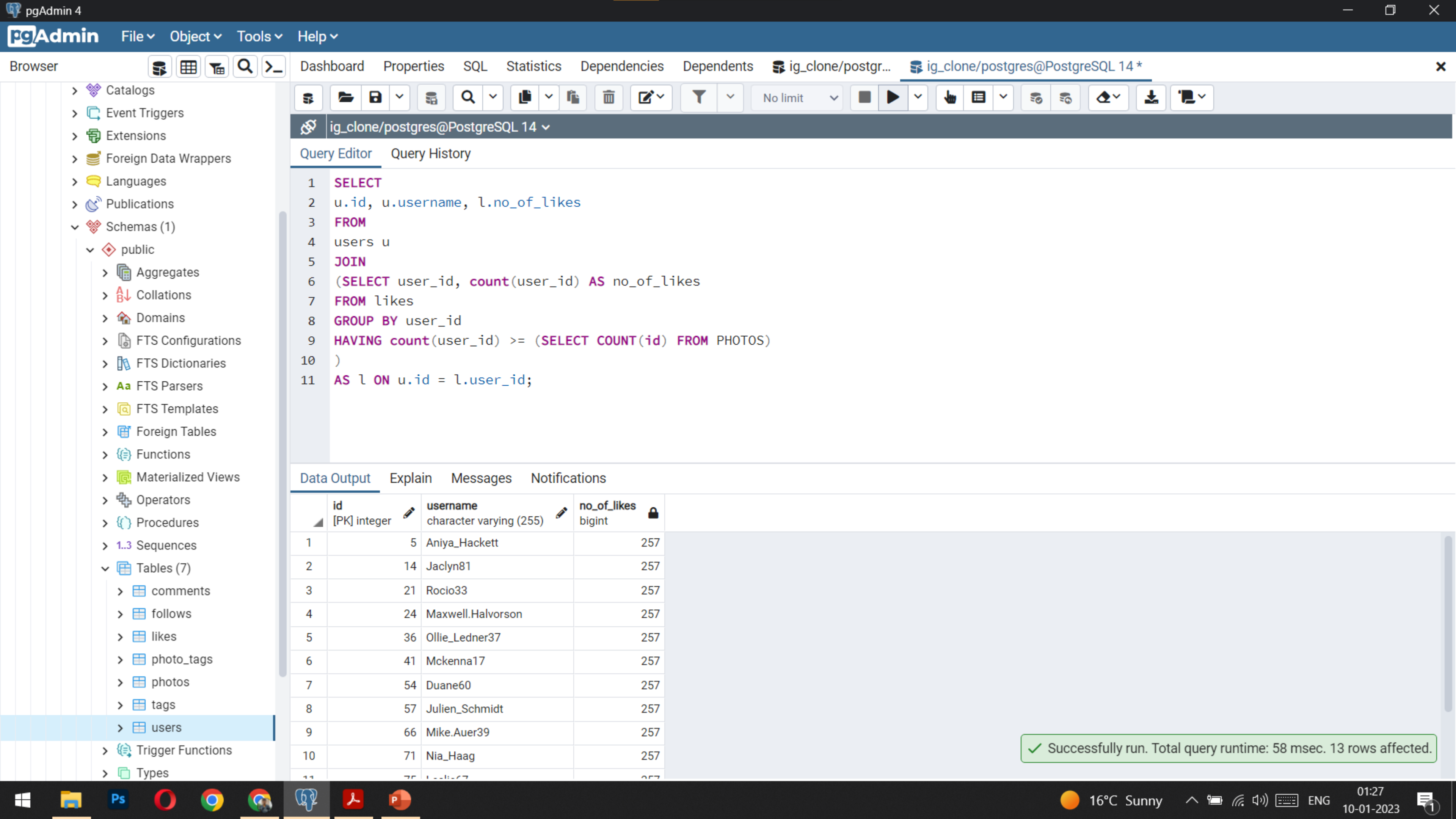
ID	User_Name	Total Post Likes
5	Aniya_Hackett	257
14	Jaclyn81	257
21	Rocio33	257
24	Maxwell.Halvorson	257
36	Ollie_Ledner37	257
41	Mckenna17	257
54	Duane60	257

ID	User_Name	Total Post Likes
57	Julien_Schmidt	257
66	Mike.Auer39	257
71	Nia_Haag	257
75	Leslie67	257
76	Janelle.Nikolaus81	257
91	Bethany20	257

2 (II) Bots & Fake Accounts

Query:-

```
SELECT
u.id, u.username, l.no_of_likes
FROM
users u
JOIN
(SELECT user_id, count(user_id) AS no_of_likes
FROM likes
GROUP BY user_id
HAVING count(user_id) >= (SELECT COUNT(id) FROM PHOTOS)
)
AS l ON u.id = l.user_id;
```



```sql
SELECT
u.id, u.username, l.no_of_likes
FROM
users u
JOIN
(SELECT user_id, count(user_id) AS no_of_likes
FROM likes
GROUP BY user_id
HAVING count(user_id) >= (SELECT COUNT(id) FROM PHOTOS)
)
AS l ON u.id = l.user_id;
```

	id [PK] integer	username character varying (255)	no_of_likes bigint
1	5	Aniya_Hackett	257
2	14	Jaclyn81	257
3	21	Rocio33	257
4	24	Maxwell.Halvorson	257
5	36	Ollie_Ledner37	257
6	41	Mckenna17	257
7	54	Duane60	257
8	57	Julien_Schmidt	257
9	66	Mike.Auer39	257
10	71	Nia_Haag	257

✓ Successfully run. Total query runtime: 58 msec. 13 rows affected.

Insights

- It took 7 days to get our first 5 users from getting our first user
- There are around 26 % user (including fake accounts) are not posted any single content yet
- There is no direct relation between posting content and getting likes (as there are some users have posted around 12 images but they not getting winning likes) “**users like Quality over Quantity**”
- Most commonly used hashtags are indicating free fun enjoying activities of users, (we can assume that people like to post their images of enjoyment)
- It is strange to that most people register on Instagram are on working days. Though Saturday weekend day still it had less no of user registration.
- There are 13 fake account who gave likes on each post of Instagram, while this account never posted any content yet.

Results & Conclusions

- This project was very helpful to understand and strengthen key concept of SQL

- It was helpful toward practicing of basic to intermediate SQL Quarries, and it was fun to write them, getting stuck and resolve error.

THANK YOU !