PROJECT - 2

Instagram User Analysis

Submitted By

Swathy Mugundan

Project 2 : Instagram User Analysis

CONTENTS

Problem Statement : Instagram User Analysis

- User analysis is the process by which we track how users engage and interact with our digital product (software or mobile application) in an attempt to derive business insights for marketing, product & development teams.
- These insights are then used by teams across the business to launch a new marketing campaign, decide on features to build for an app, track the success of the app by measuring user engagement and improve the experience altogether while helping the business grow.
- You are working with the product team of Instagram and the product manager has asked you to provide insights on the questions asked by the management team.

Project Description : Marketing Analysis

01

Rewarding Most Loyal Users

Finding the 5 oldest users who have been using the platform for longest time

02

Remind Inactive Users to start posting

Finding users who have never posted a single photo on Instagram

03

Declaring Contest Winners

User who gets most likes on a single photo is declared as winner and details are fetched

04

Hasgtag Re-Searching

Identify the most-used hashtags on Instagram and help partner brand to reach out to most people

05

Launch Ad-Campaign

Identify which day would be the best to launch an ad-campaign by choosing which day most registrations had happen

Project Description : Investor Metrics

User Engagement

Find the activity status of the users by checking their average posts, total number of photos and total number of users

Bots & Fake Accounts

Find the users who have liked all the posts (possibly a bot) :

This helps to find the fake/dummy accounts on the platform

Project Approach : Instagram Analysis

DB Data

Download the dataset from the site, run the queries one by one and form the database with all the required tables.

Problem Def'n

Analyse the tables and its attributes and understand the references between tables. Read the question properly.

Query Formation

Figure out the related tables and write the suitable SQL Statement for fetching the results

Result Analysis

Analyse and understand the result set and capture the inference of the problem statement,

Tech Stack Used

My SQL Server

User for the Relational Database Management System. To create database, To create Tables.

MySQL Workbench

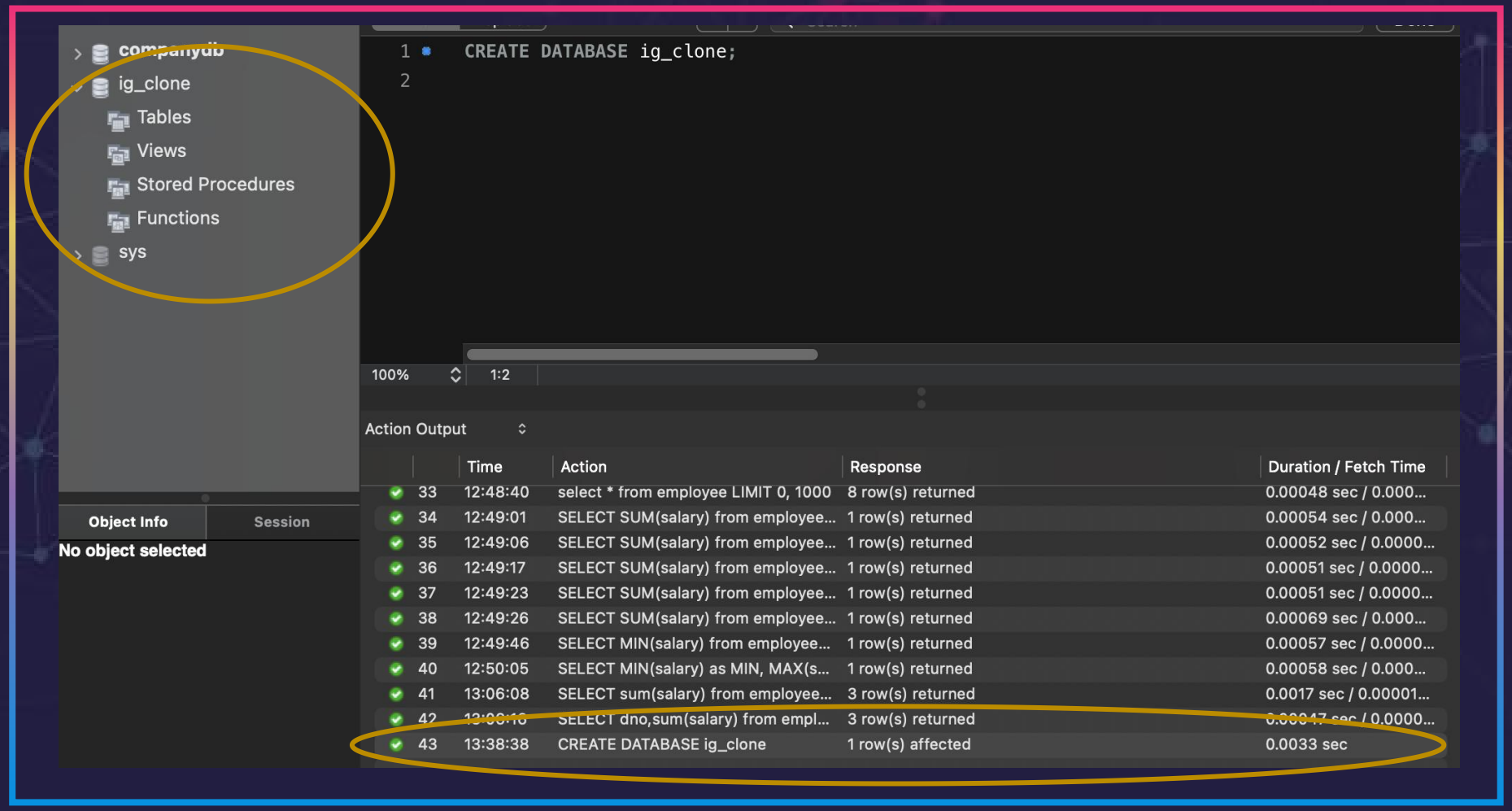
It is a GUI Based tool for Querying the Database. It is a visual-based DB tool to observe the databases , its tables, its attributes and how they are being stored in the system.

Teampaper Snap

To take screenshots of the result-grid of the SQL Statements from SQL Workbench

Project 2 : Instagram User Analysis

Database Creation Screenshots



Project 2 : DB-Tables Creation Screenshots

Users Table

```sql
/*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'), ('Ha

select * from users;
```

100% 1:9

Result Grid | Filter Rows: Search | Edit: | Export/Import:

id	username	created_at
1	Kenton_Kirlin	2017-02-16 18:22:11
2	Andre_Purdy85	2017-04-02 17:11:21
3	Harley_Lind18	2017-02-21 11:12:33
4	Arely_Bogan63	2016-08-13 01:28:43
5	Aniya_Hackett	2016-12-07 01:04:39
6	Travon.Waters	2017-04-30 13:26:14
7	Kasandra_Homenick	2016-12-12 06:50:08
8	Tabitha_Schamberger11	2016-08-20 02:19:46
9	Gus93	2016-06-24 19:36:31
10	Presley_McClure	2016-08-07 16:25:49
11	Justina.Gaylord27	2017-05-04 16:32:16
12	Dereck65	2017-01-19 01:34:14
13	Alexandro35	2017-03-29 17:09:02
14	Jaclyn81	2017-02-06 23:29:16

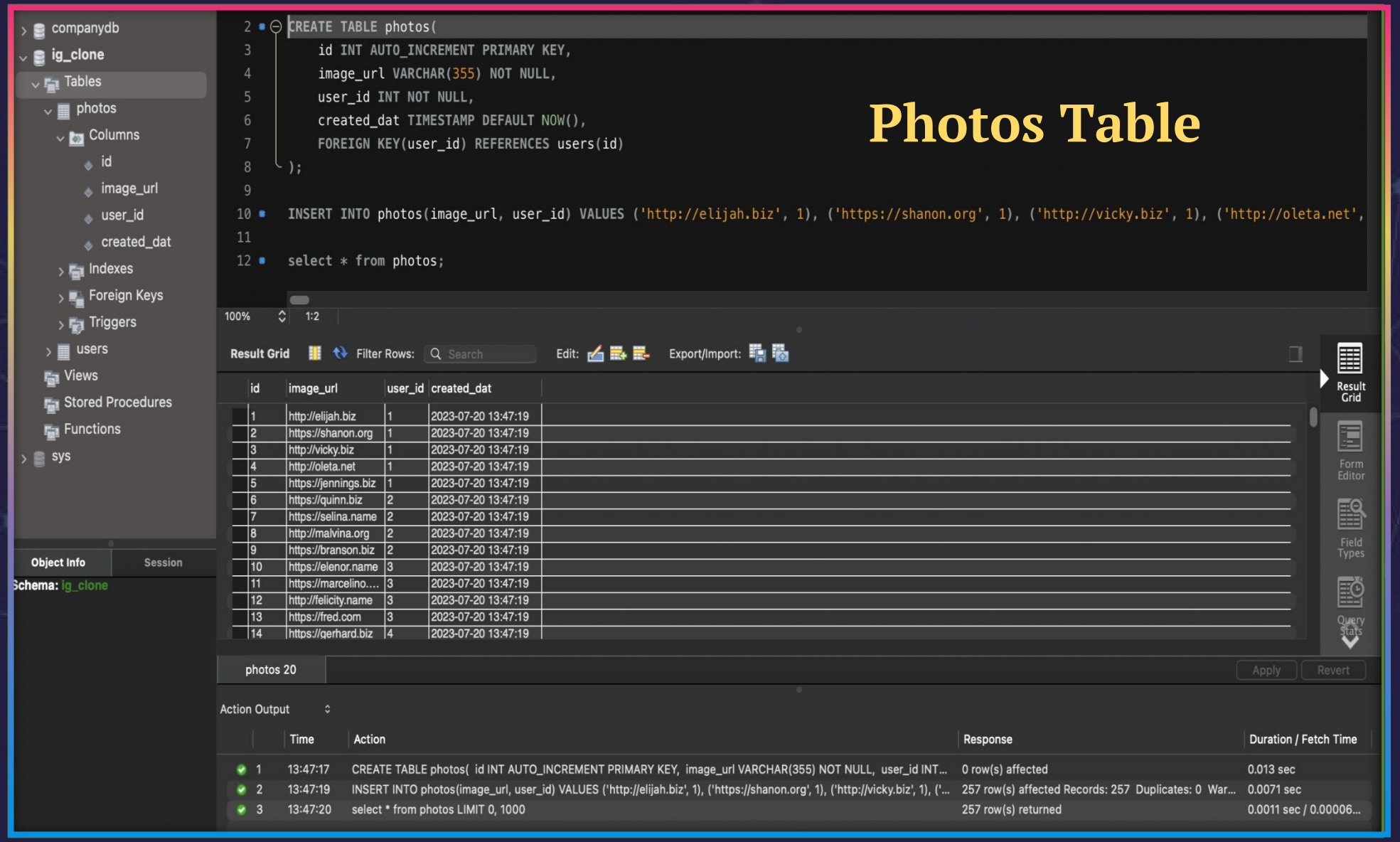
Result Grid | Form Editor | Field Types | Query Stats

users 19 | Apply | Revert

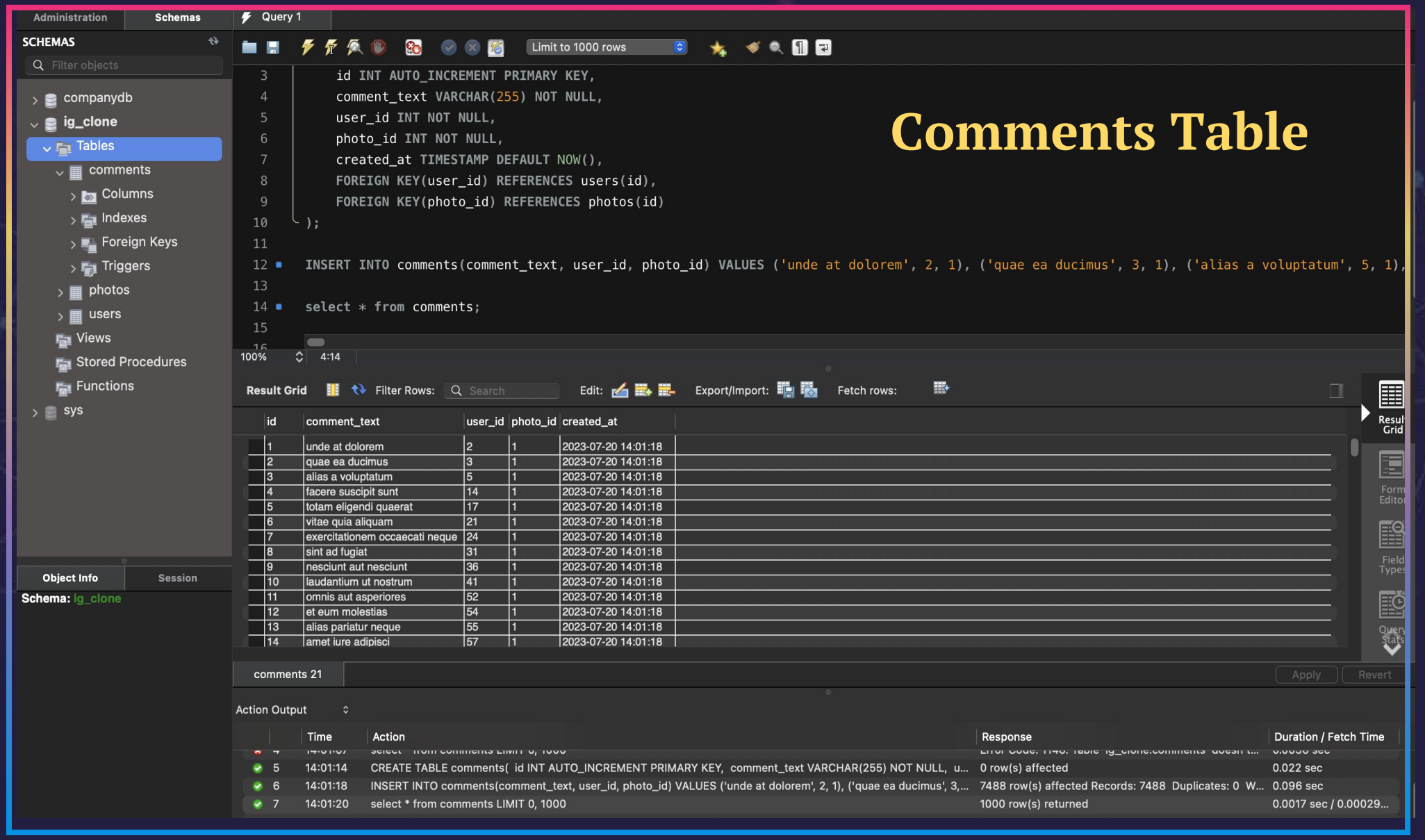
Action Output

	Time	Action	Response	Duration / Fetch Time
1	13:42:24	CREATE TABLE users(id INT AUTO_INCREMENT UNIQUE PRIMARY KEY, username VARCHAR(255) NOT NULL, creat...	0 row(s) affected	0.011 sec
2	13:42:48	INSERT INTO users (username, created_at) VALUES ('Kenton_Kirlin', '2017-02-16 18:22:10.846'), ('Andre_Purdy85', '2...	100 row(s) affected Records: 100 Duplicates: 0 War...	0.0054 sec
3	13:43:06	select * from users LIMIT 0, 1000	100 row(s) returned	0.00080 sec / 0.000...

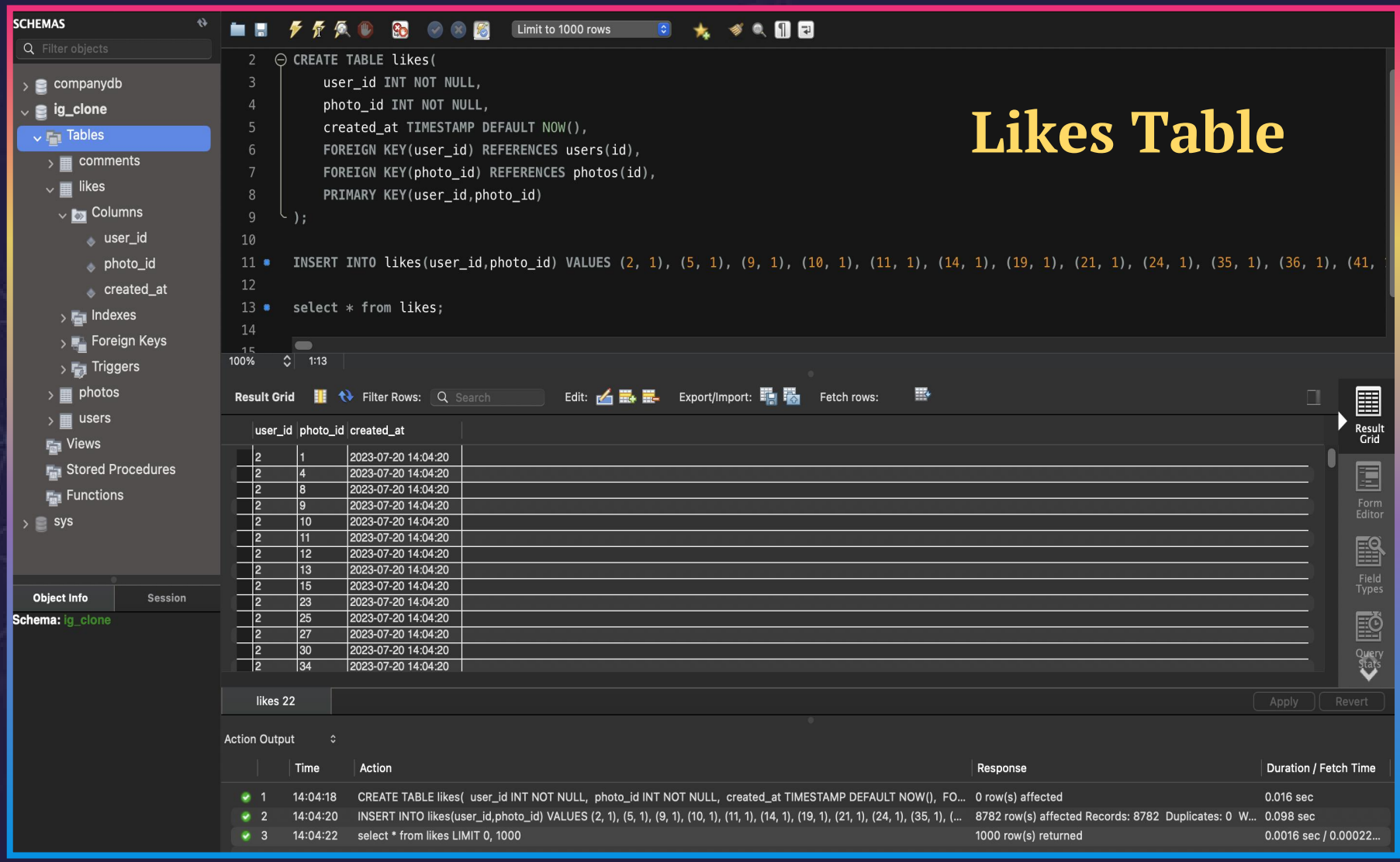
Project 2 : DB-Tables Creation Screenshots



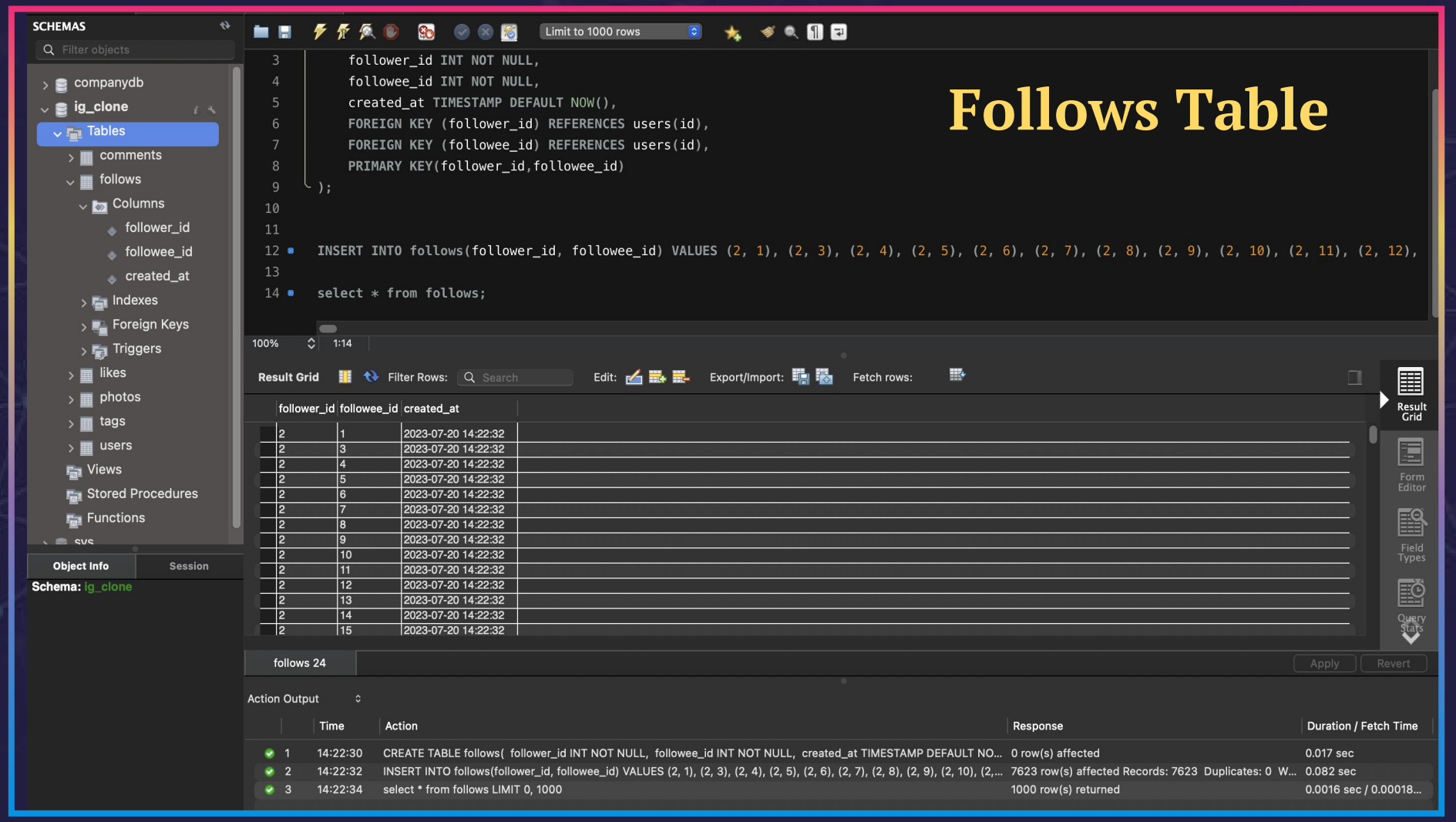
Project 2 : DB-Tables Creation Screenshots



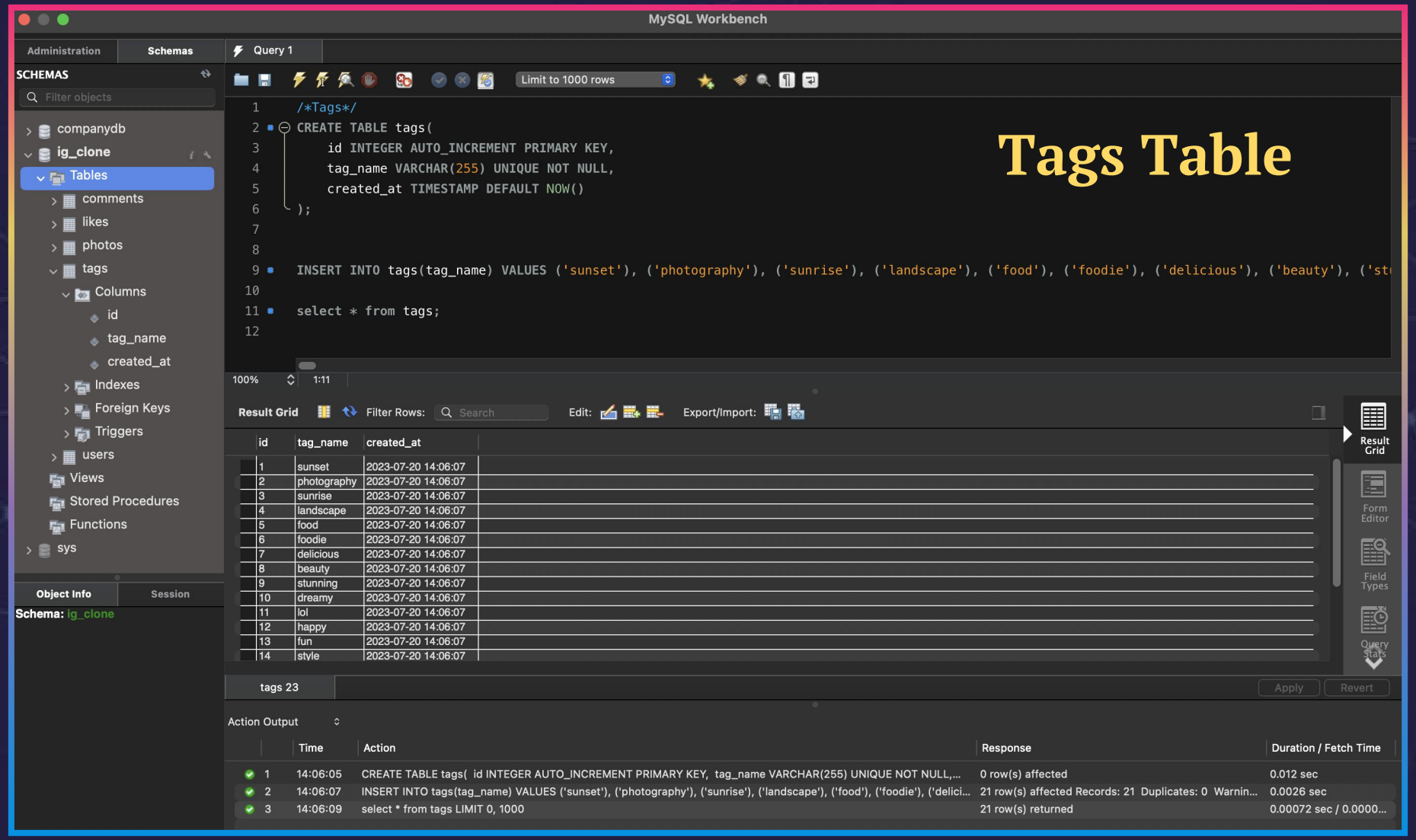
Project 2 : DB-Tables Creation Screenshots



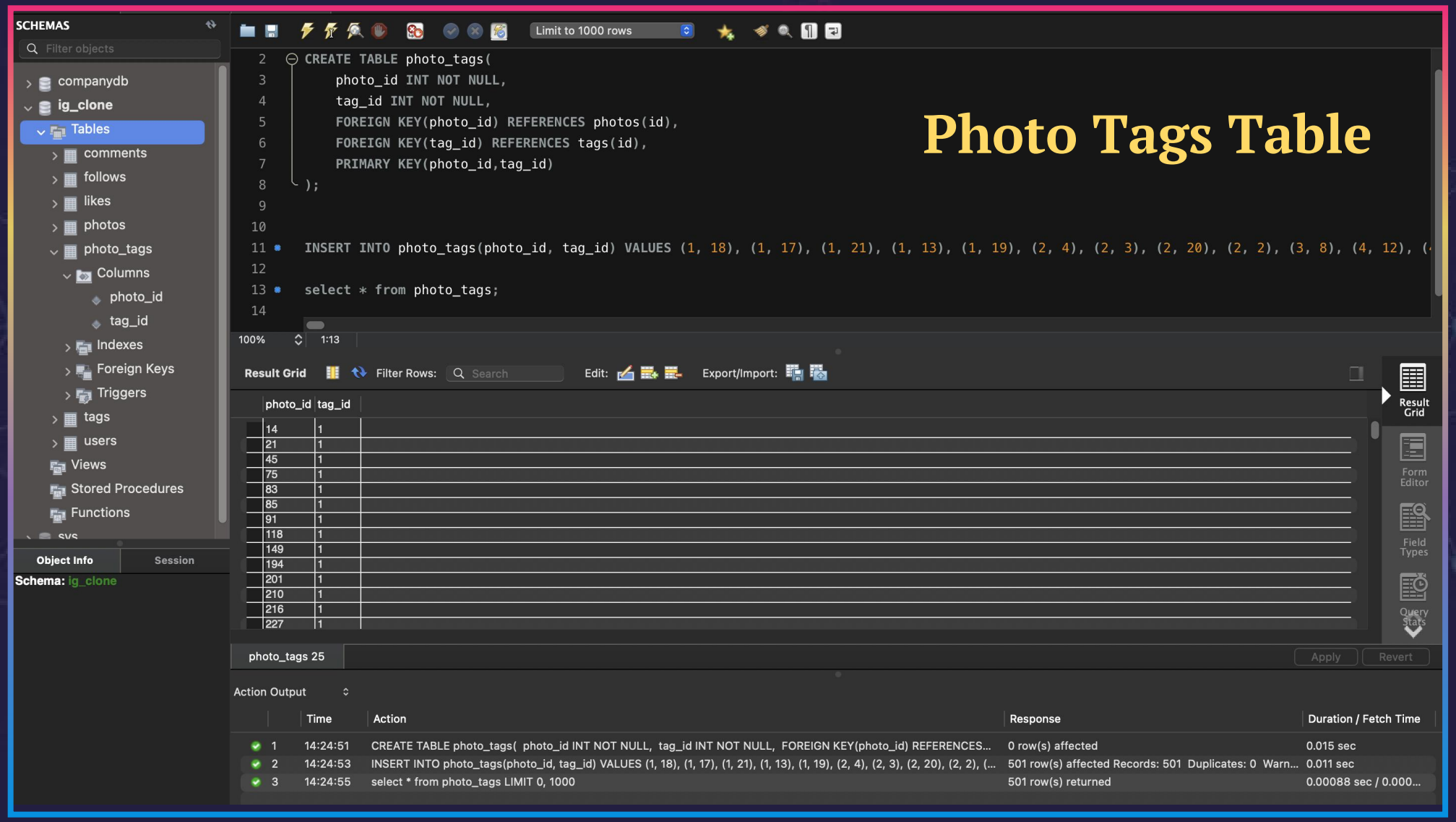
Project 2 : DB-Tables Creation Screenshots



Project 2 : DB-Tables Creation Screenshots



Project 2 : DB-Tables Creation Screenshots



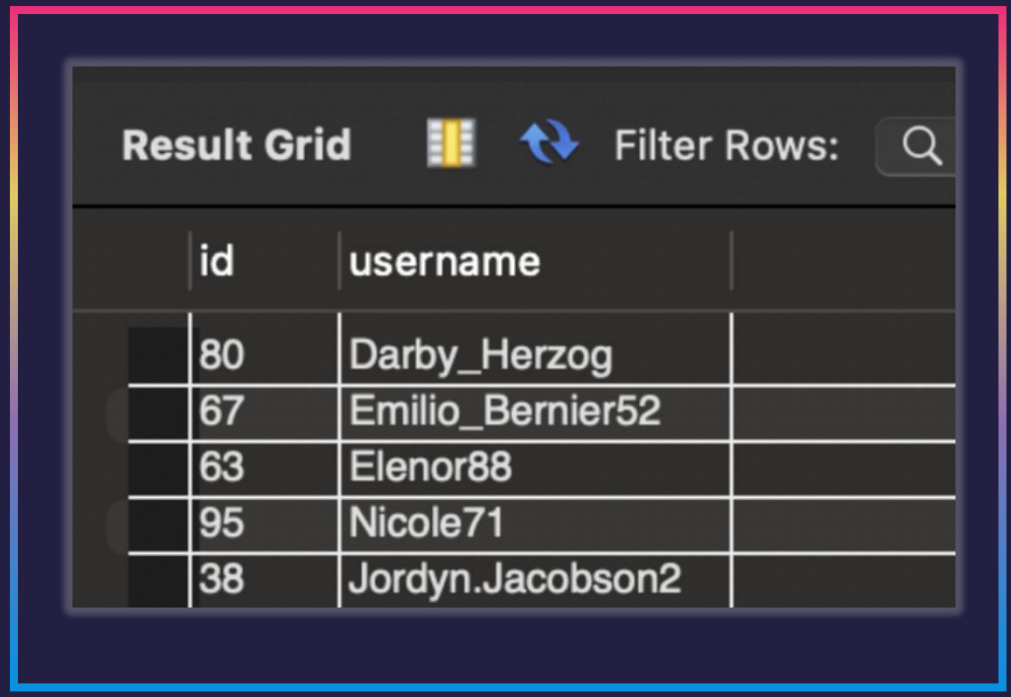
Marketing Analysis : Loyal User Reward

```
select   id,  username

from users

order by created_at

LIMIT 5;
```



Result Grid Filter Rows:

id	username
80	Darby_Herzog
67	Emilio_Bernier52
63	Elenor88
95	Nicole71
38	Jordyn.Jacobson2

Marketing Analysis : Inactive Users (Zero Photos Posted)

```
select users.id, username

    from users left join photos

    on users.id = photos.user_id

where image_url is null;
```

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
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
74	Hulda.Macejkovic
75	Leslie67
76	Janelle.Nikolaus81
80	Darby_Herzog
81	Esther.Zulauf61
83	Bartholome.Bernhard
89	Jessyca_West
90	Esmeralda.Mraz57
91	Bethany20

Marketing Analysis : Contest Winner Declaration

```
select U.id, U.username, P.photo_id, P.count as Likes
from users U
join
(select photos.user_id, photo_id, count(photo_id) as count
        from likes left join photos
        on likes.photo_id = photos.id
        group by likes.photo_id
        order by count desc) P
on U.id = P.user_id
order by likes desc;
```

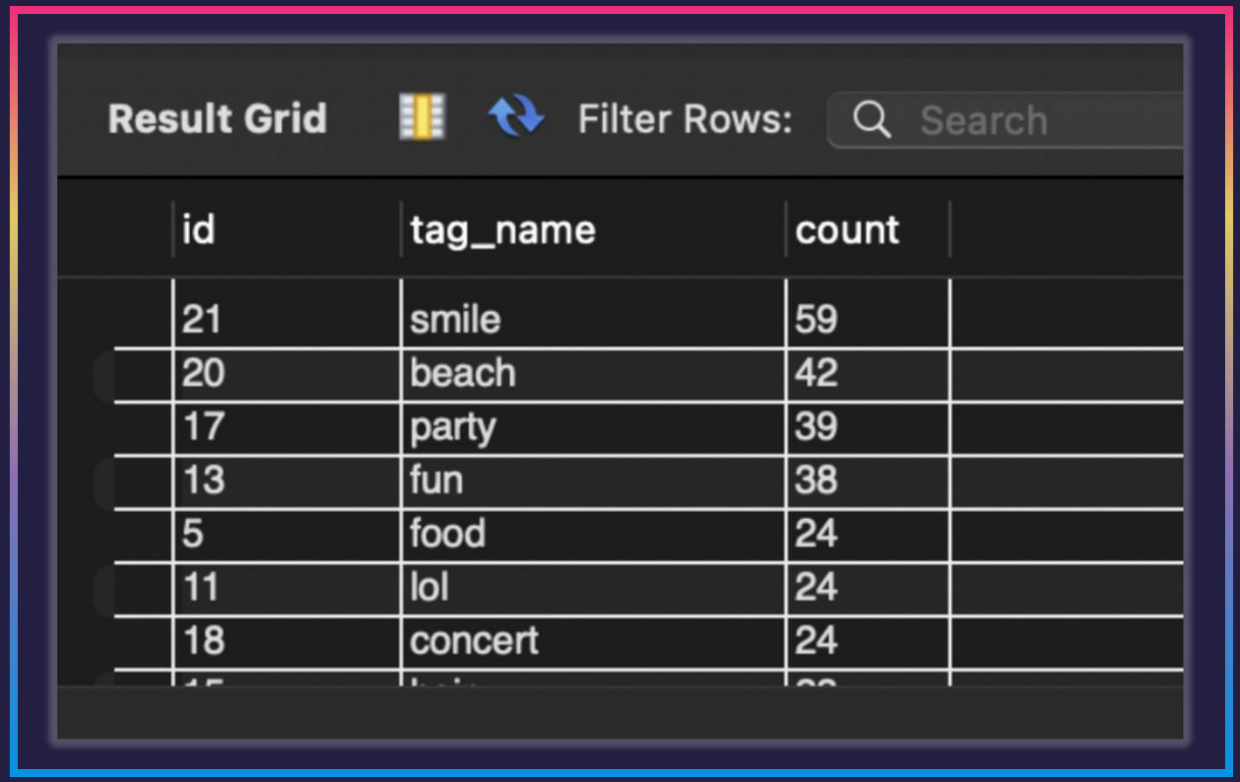
Result Grid | Filter Rows: Search

id	username	photo_id	Likes
52	Zack_Kemmer93	145	48
46	Malinda_Streich	127	43
65	Adelle96	182	43
44	Seth46	123	42
10	Presley_McClure	30	41
16	Annalise.McKenzie16	52	41
20	Delpha.Kihn	61	41
55	Meggie_Doyle	147	41
63	Elenor88	174	41
72	Kathryn80	192	41
100	Javonte83	256	41
3	Harley_Lind18	13	40
32	Irwin.Larson	97	40
58	Aurelie71	153	40
59	Cesar93	161	40
94	Damon35	244	40
13	Alexandro35	44	39
22	Kenneth64	62	39
23	Eveline95	66	39
33	Yvette.Gottlieb91	100	39
37	Yazmin_Mills95	107	39

CONCLUSION : Based on the result set, we can conclude that Zack_Kemmer93 is the winner for securing most likes in single photo

Marketing Analysis : Hastag Research

```
select id, tag_name, count(tag_name) as count
from tags left join photo_tags
on tags.id = photo_tags.tag_id
group by id
order by count desc ;
```



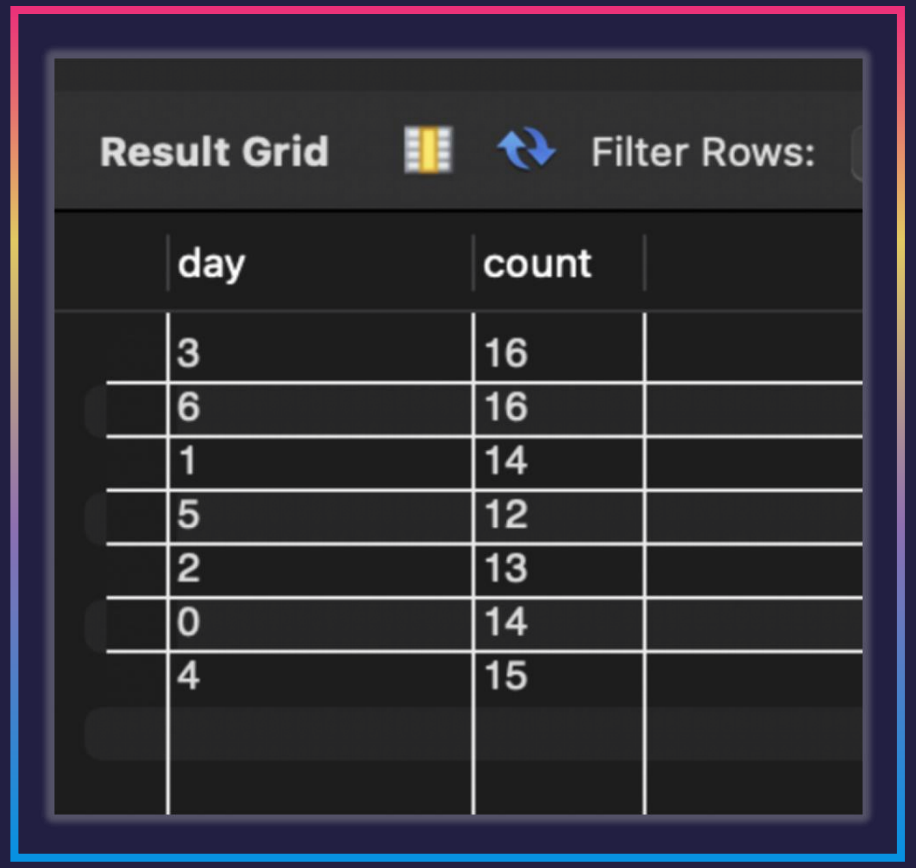
id	tag_name	count
21	smile	59
20	beach	42
17	party	39
13	fun	38
5	food	24
11	lol	24
18	concert	24

CONCLUSION

Based on the result set, we can conclude that most of the users used the “SMILE” hashtag.

Marketing Analysis : Ad-Campaign

```
select  weekday(created_at) as day,
count(weekday(created_at)) as count
from users
group by day;
```



day	count
3	16
6	16
1	14
5	12
2	13
0	14
4	15

CONCLUSION

Based on the result set, we can conclude that most of the users have registered to the platform on Wednesday (3) and Saturday (6). Hence, ad campaign can be planned on those days.

Investor Metrics : Average Analysis (1)

```
select user_id as ID,
count(user_id) as No_of_Posts,
257/count(user_id) as Average
from photos
group by user_id;
```

NOTE

The total number of posts is 257 and users is 100.

Result Grid Filter Rows: Search

ID	No_of_Posts	Average
1	5	51.4000
2	4	64.2500
3	4	64.2500
4	3	85.6667
6	5	51.4000
8	4	64.2500
9	4	64.2500
10	3	85.6667
11	5	51.4000
12	4	64.2500
13	5	51.4000
15	4	64.2500
16	4	64.2500
17	3	85.6667
18	1	257.0000
19	2	128.5000
20	1	257.0000
22	1	257.0000
23	12	21.4167
26	5	51.4000

Result Grid Filter Rows: Sear

ID	No_of_Posts	Average
27	1	257.0000
28	4	64.2500
29	8	32.1250
30	2	128.5000
31	1	257.0000
32	4	64.2500
33	5	51.4000
35	2	128.5000
37	1	257.0000
38	2	128.5000
39	1	257.0000
40	1	257.0000
42	3	85.6667
43	5	51.4000
44	4	64.2500
46	4	64.2500
47	5	51.4000
48	1	257.0000
50	3	85.6667
51	5	51.4000

Investor Metrics : Average Analysis (2)

```
select user_id as ID,
count(user_id) as No_of_Posts,
257/count(user_id) as Average
from photos
group by user_id;
```

NOTE

The total number of posts is 257 and users is 100.

Result Grid Filter Rows: Search

ID	No_of_Posts	Average
52	5	51.4000
55	1	257.0000
56	1	257.0000
58	8	32.1250
59	10	25.7000
60	2	128.5000
61	1	257.0000
62	2	128.5000
63	4	64.2500
64	5	51.4000
65	5	51.4000
67	3	85.6667
69	1	257.0000
70	1	257.0000
72	5	51.4000

Result Grid Filter Rows:

ID	No_of_Posts	Average
72	5	51.4000
73	1	257.0000
77	6	42.8333
78	5	51.4000
79	1	257.0000
82	2	128.5000
84	2	128.5000
85	2	128.5000
86	9	28.5556
87	4	64.2500
88	11	23.3636
92	3	85.6667
93	2	128.5000
94	1	257.0000
95	2	128.5000
96	3	85.6667
97	2	128.5000
98	1	257.0000
99	3	85.6667
100	2	128.5000

Investor Metrics : Bots & Fake Accounts

```
select U.id as ID, U.username as Uname, L.count as Likes
from users U      join
(select user_id, count(user_id) as count from likes
            group by user_id
            order by count desc)  L
on U.id = L.user_id
where L.count = 257
group by U.id;
```

NOTE

The total number of posts is 257.

If a user is liking all 257 posts

then it can possibly be a bot.

Result Grid Filter Rows:

ID	Uname	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
57	Julien_Schmidt	257
66	Mike.Auer39	257
71	Nia_Haag	257
75	Leslie67	257
76	Janelle.Nikolaus81	257
91	Bethany20	257

Project 2 : Result Analysis & Insights

- The required database is formed and the tables are populated. SQL Statements were ran on the DB to fetch the results of the required problem statements.
- Marketing Analysis steps helps us to understand the users and their respective actions on the Instagram platform.
- Investor Analysis helps us to observe the average user posts and the fake bots available with the instagram platform
- Thus, using a simple SQL Query we were able to analyse the data and figured out certain metrics and analytics upon it.

Thank You

Submitted By

Swathy Mugundan