INSTAGRAM ANALYTICS

PROJECT DESCRIPTION- This project is based on user interaction and engagement with the Instagram social networking site. Based on the data provided we are supposed to get useful insights that can help the company grow. Data comprised of 7 different tables that contain information about the user's login, photo uploads, tags used, comments etc.

The main purpose of this project is to provide useful insights that can help managers make the right decisions for the company's betterment.

APPROACH= My analysis started by studying the data properly and thoroughly which includes understanding which table contains what information and details about primary key and foreign key which later helped me to answer the questions.

SOFTWARE USED- SQL Workbench

Question- 1 Identify the five oldest users on Instagram from the provided database.

```
1 •    select * from users
2      order by created_at
3      limit 5;
```

Result Grid | Filter Rows: | Edit:

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.Jacobson2	2016-05-14 07:56:26

users 8 ×

Quesiton: Identify users who have never posted a single photo on Instagram.

```
1   select a.username,a.id from users a
2   left join photos b
3   on a.id=b.user_id
4   where b.image_url is null
5   order by a.username
```

Result Grid | Filter Rows: | Export

username	id
Aniya_Hackett	5
Bartholome.Bernhard	83
Bethany20	91
Darby_Herzog	80
David.Osinski47	45
Duane60	54
Esmeralda.Mraz57	90
Esther.Zulauf61	81
Franco_Keebler64	68
Hulda.Macejkovic	74
Jaclyn81	14
Janelle.Nikolaus81	76
Jessyca West	89

Result 21 ×

Question: Determine the winner of the contest and provide their details to the team.(Max Likes)

```
1 •  select a.photo_id,b.username,count(a.user_id)as number_of_likes from likes a
2     inner join photos c
3     on a.photo_id=c.id
4     inner join users b
5     on c.user_id=b.id
6     group by a.photo_id,b.username
7     order by number_of_likes desc;
```
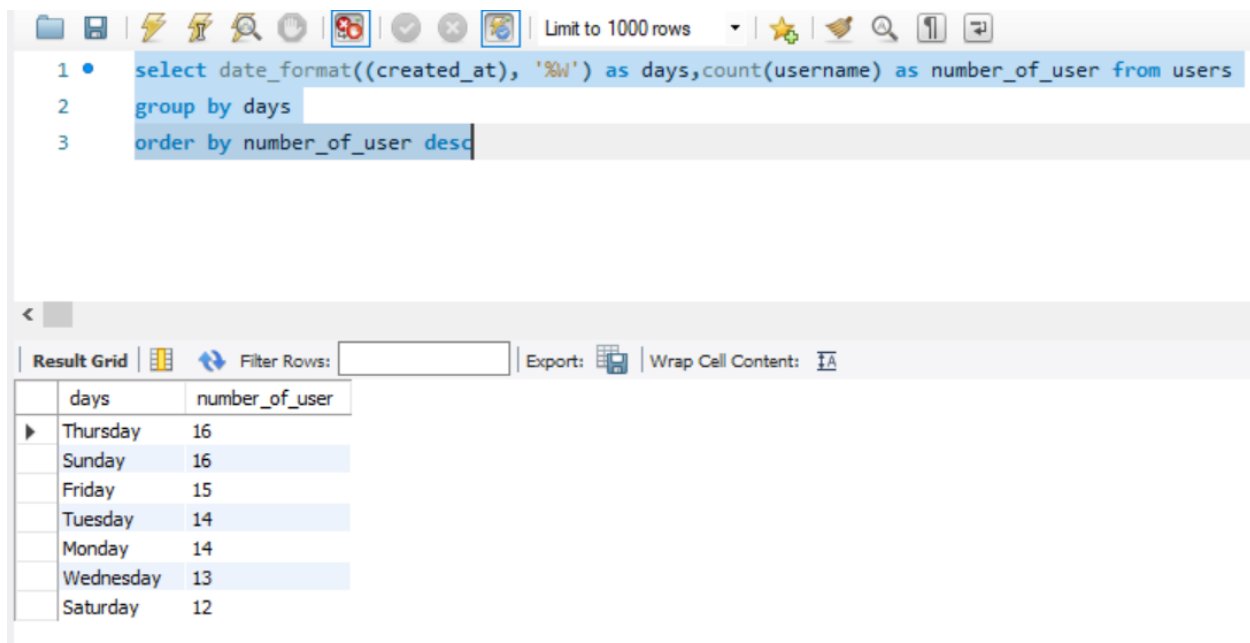
Result Grid | Filter Rows: | Export: | Wrap Cell Content: ᴵᴬ

photo_id	username	number_of_likes
145	Zack_Kemmer93	48
127	Malinda_Streich	43
182	Adelle96	43
123	Seth46	42
30	Presley_McClure	41
52	Annalise.McKenzie16	41
61	Delpha.Kihn	41
147	Meggie_Doyle	41
174	Elenor88	41
192	Kathryn80	41
256	Javonte83	41
13	Harley_Lind18	40
97	Irwin.Larson	40
153	Aurelie71	40

Question: Identify and suggest the top five most commonly used hashtags on the platform.

```
1    select a.tag_name,count(b.photo_id) as numbers from photo_tags b
2    inner join tags a
3    on a.id=b.tag_id
4    group by a.tag_name
5    order by numbers desc
6    limit 5;
7
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: $\overline{I}A$ | Fetch rows:

tag_name	numbers
smile	59
beach	42
party	39
fun	38
concert	24

Your Task: Determine the day of the week when most users register on Instagram. Provide insights on when to schedule an ad campaign.
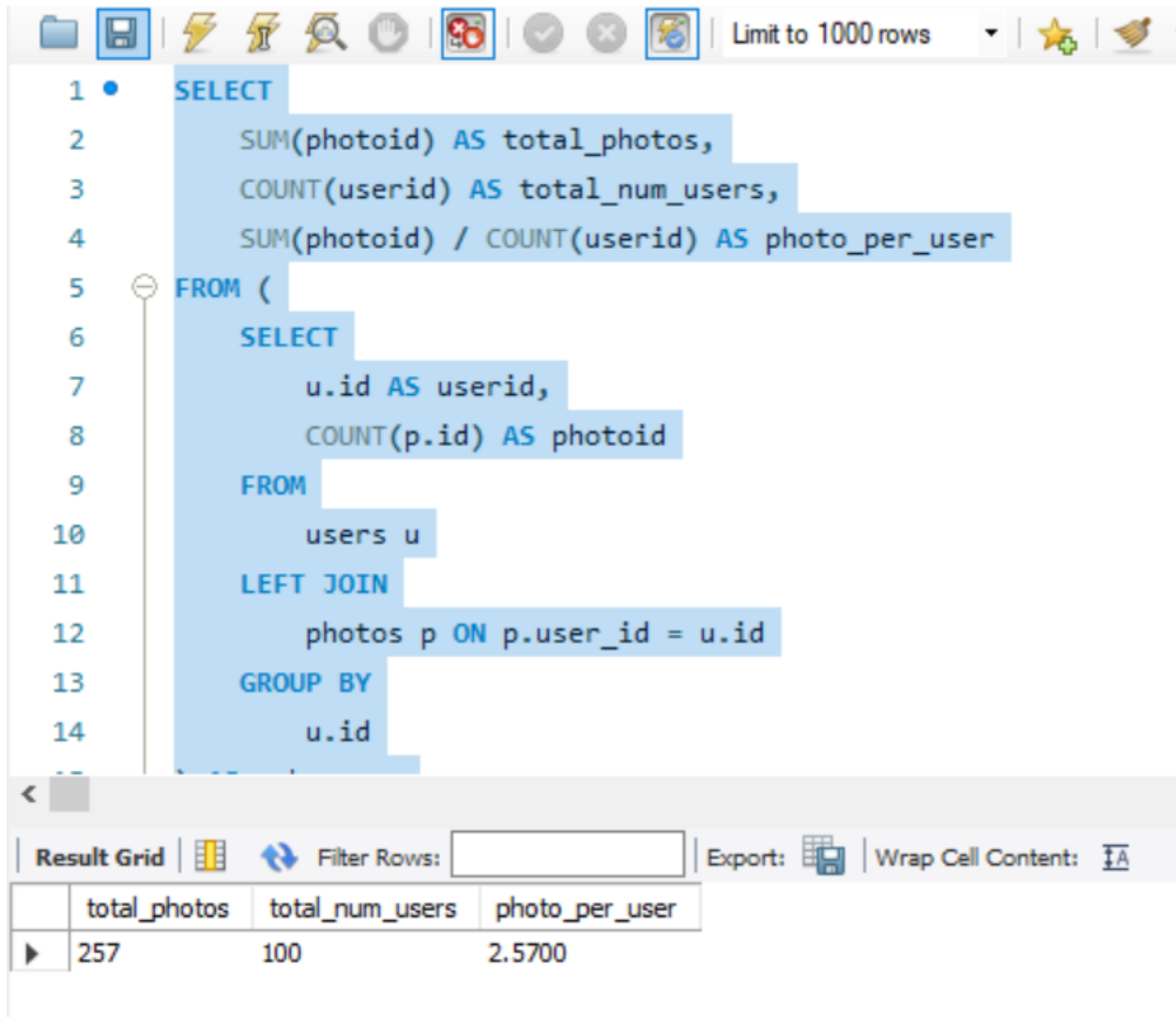
```sql
1 ● select date_format((created_at), '%W') as days,count(username) as number_of_user from users
2   group by days
3   order by number_of_user desc
```

Limit to 1000 rows

Result Grid | Filter Rows: | Export: | Wrap Cell Content: $\overline{I}A$

days	number_of_user
Thursday	16
Sunday	16
Friday	15
Tuesday	14
Monday	14
Wednesday	13
Saturday	12

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.

```sql
 1 ●    SELECT
 2          SUM(photoid) AS total_photos,
 3          COUNT(userid) AS total_num_users,
 4          SUM(photoid) / COUNT(userid) AS photo_per_user
 5      FROM (
 6          SELECT
 7              u.id AS userid,
 8              COUNT(p.id) AS photoid
 9          FROM
10              users u
11          LEFT JOIN
12              photos p ON p.user_id = u.id
13          GROUP BY
14              u.id
```

	total_photos	total_num_users	photo_per_user
▶	257	100	2.5700

Identify users (potential bots) who have liked every single photo on the site, as this is not typically possible for a normal user.

```
 1 •   SELECT username
 2     FROM users
 3   ⊖ WHERE id IN (SELECT DISTINCT user_id
 4                  FROM Likes
 5                  WHERE user_id IS NOT NULL
 6                  GROUP BY user_id
 7   ⊖             HAVING COUNT(DISTINCT photo_id) = (SELECT
 8                          COUNT(DISTINCT photo_id)
 9                      FROM
10                          Likes)
11                  ORDER BY username);
```

username
Aniya_Hackett
Jaclyn81
Rocio33
Maxwell.Halvorson
Ollie_Ledner37
Mckenna17
Duane60
Julien_Schmidt
Mike.Auer39
Nia_Haag

users 16 ✕

Insights-

    a. We have data of 100 users.
    b. Total photos uploaded 257.
    c. First user who signed up was Darby Herzog.
    d. Out of all the photos uploaded Zack kemmer got most likes on his photo.
    e. Smile is most used hashtag.
    f. Mostly people registered on Thursday

I learnt and was able to apply my learnings to solve the queries. I got my doubts cleared about sub queries and joins and now I am in a much better position to use them as and when it is required.