INSTAGRAM USER ANALYTICS

SAMIYA ALAM

29/05/2024

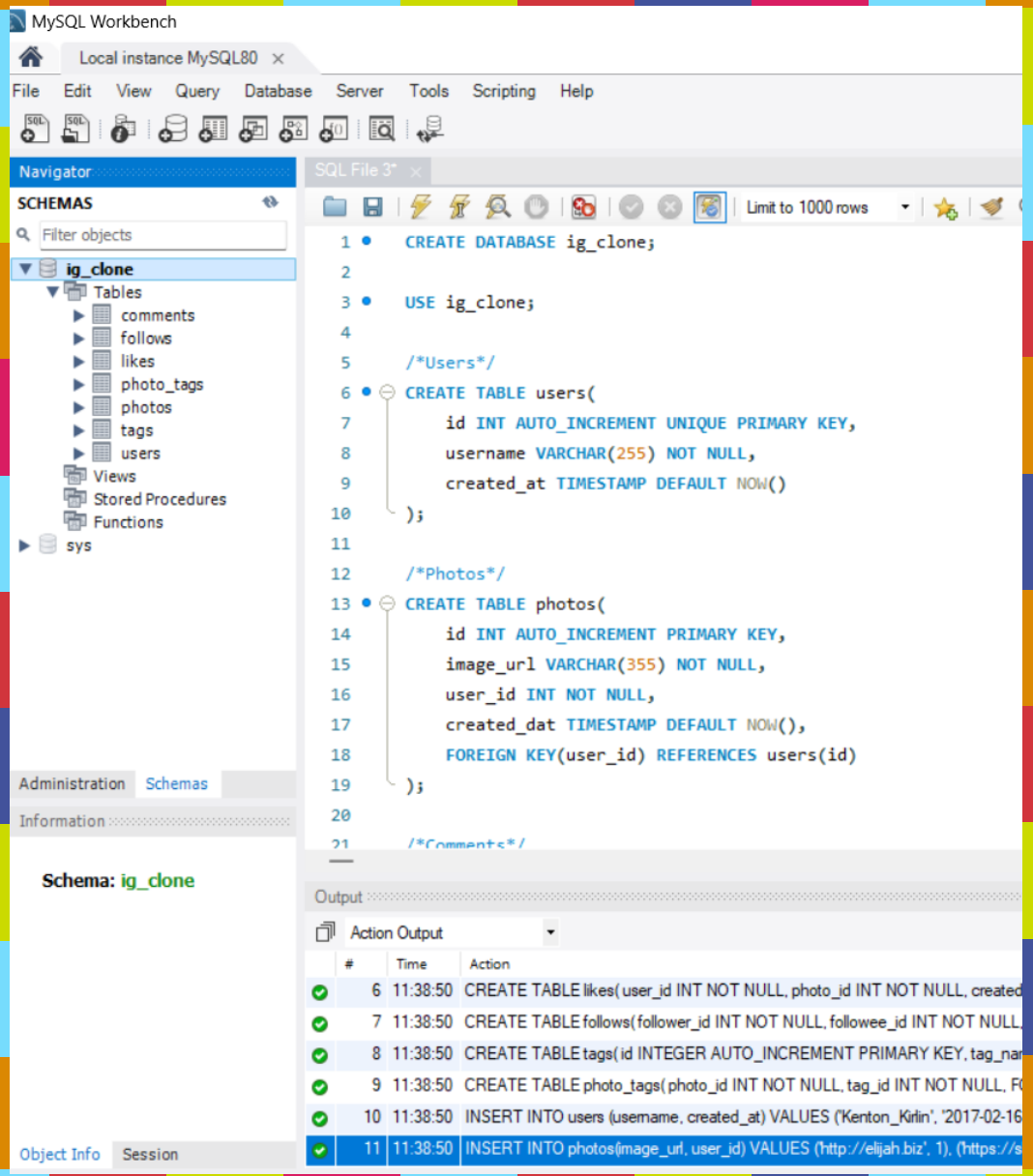
Project Description

This project involves detailed SQL queries and data manipulation to answer key business questions for the Instagram platform. The insights gained will help the marketing team with campaigns, the product team with feature development, and provide investors with important metrics on user engagement and platform integrity. Through systematic data analysis, we will support the growth and enhancement of Instagram's user experience and business strategy.

TECH STACK USED IN THIS PROJECT :

- MySQL Workbench for writing queries
- Microsoft PowerPoint for making the presentation

INSIGHTS WITH QUERIES, OUTPUT AND APPROACH



Database Creation:

A DATABASE named ig_clone is created using queries given in project description

Marketing Analysis

Loyal User Reward:

Task: Identify the five oldest users on Instagram.
Approach: Query the database to find the users with the earliest registration dates and list the top five.

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

QUERY

Result Grid | Filter Rows:

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
NULL	NULL	NULL

OUTPUT

Inactive User Engagement:

Task: Identify users who have never posted a single photo.

Approach: Filter the user data to find those with zero photo posts and compile a list for promotional outreach.

```
SELECT users.id, users.username
FROM users
LEFT OUTER JOIN photos ON users.id=photos.user_id
WHERE photos.user_id IS NULL;
```

QUERY

id	username
5	Aniya_Hackett
7	Kasandra_Homenick
14	Jadyn81
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
75	Leslie67
76	Janelle.Nikolaus81
80	Darby_Herzog
81	Esther.Zulauf61
83	Bartholome.Bernhard
89	Jessyca_West
90	Esmeralda.Mraz57
91	Bethany20

OUTPUT

Contest Winner Declaration:

Task: Determine the user with the most likes on a single photo.
Approach: Use data to identify the photo with the highest number of likes and provide the corresponding user details.

WINNER: Zack_kemmer93

```sql
SELECT users.username,  photos.id, photos.image_url, COUNT(*) AS Number_of_likes
FROM likes
JOIN photos ON photos.id = likes.photo_id
JOIN users ON users.id = photos.user_id
GROUP BY photos.id
ORDER BY Number_of_likes DESC
LIMIT 10;
```

Result Grid | Filter Rows: | Export: | Wrap C

username	id	image_url	Number_of_likes
Zack_Kemmer93	145	https://jarret.name	48
Adelle96	182	https://dorcas.biz	43
Malinda_Streich	127	https://celestine.name	43
Seth46	123	http://shannon.org	42
Presley_McClure	30	http://kenny.com	41
Elenor88	174	https://delbert.net	41
Kathryn80	192	https://anahi.info	41
Meggie_Doyle	147	https://adela.com	41
Delpha.Kihn	61	https://dejon.name	41
Annalise.McKenzie16	52	https://hershel.com	41

QUERY

OUTPUT

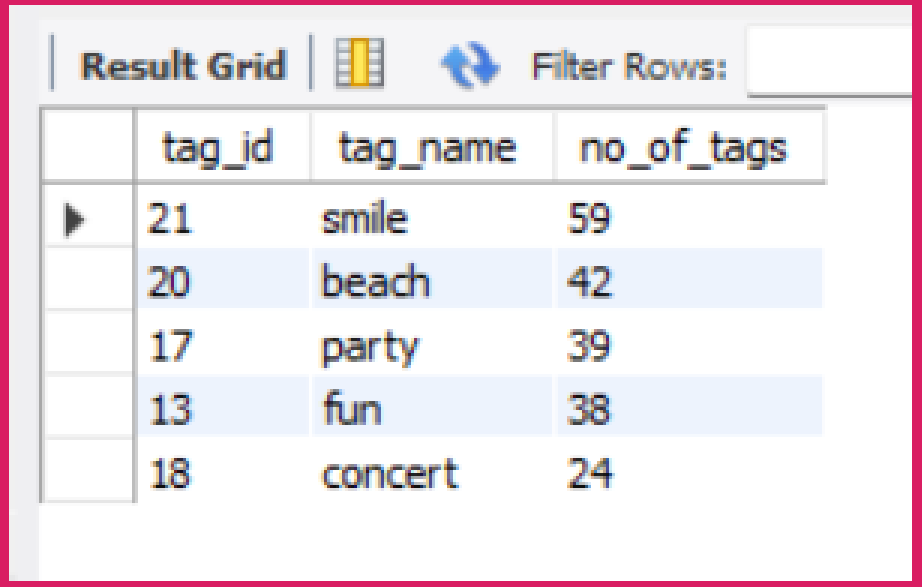
Hashtag Research:

Task: Identify the top five most commonly used hashtags.

Approach: Analyze the hashtag data to count occurrences and rank the top five hashtags.

```
select tag_id, tag_name, count(*) as no_of_tags
from photo_tags
inner join tags on
tags.id=photo_tags.tag_id
group by tag_id
order by no_of_tags
DESC
limit 5;
```

QUERY



tag_id	tag_name	no_of_tags
21	smile	59
20	beach	42
17	party	39
13	fun	38
18	concert	24

OUTPUT

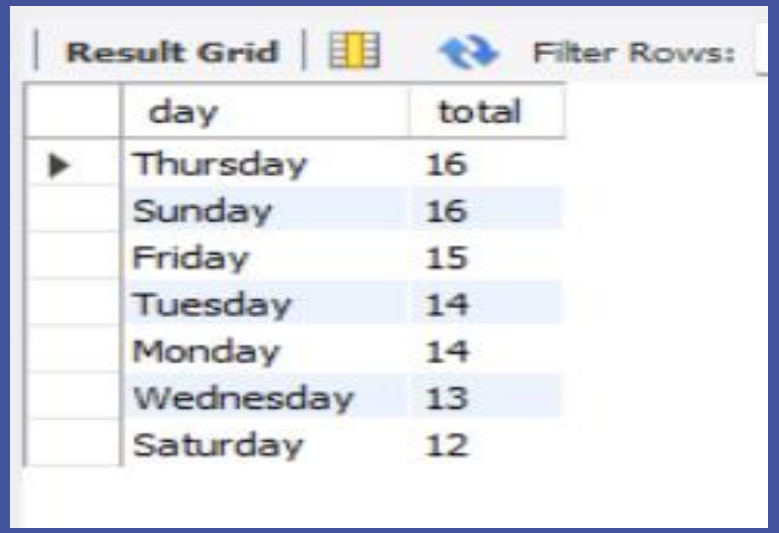
Ad Campaign Launch:

Task: Determine the best day of the week to launch ads based on user registrations.
Approach: Examine the data by day of the week to find the day with the highest number of user sign-ups.

```
SELECT DAYNAME(created_at) as day, count(*) as total from users
group by day
order by total desc;
```

QUERY

SUNDAY & THURSDAY



day	total
Thursday	16
Sunday	16
Friday	15
Tuesday	14
Monday	14
Wednesday	13
Saturday	12

OUTPUT

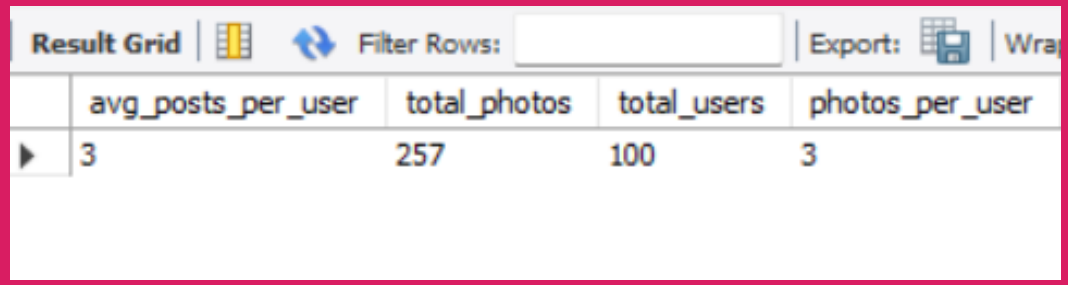
Investor Metrics

User Engagement:

Task: Calculate the average number of posts per user and provide the total number of photos divided by the total number of users.

Approach: Compute the average by dividing the total number of posts by the total number of users, and then provide the relevant metrics

```sql
SELECT ceil(COUNT(photos.id) / COUNT(DISTINCT users.id)) AS avg_posts_per_user,
COUNT(photos.id) AS total_photos, COUNT(DISTINCT users.id) AS total_users,
CEIL( COUNT(photos.id) / COUNT(DISTINCT users.id)) AS photos_per_user
FROM users
LEFT JOIN photos ON users.id = photos.user_id;
```



avg_posts_per_user	total_photos	total_users	photos_per_user
3	257	100	3

QUERY

OUTPUT

Bots & Fake Accounts:

Task: Identify users who have liked every single photo, indicative of potential bot behavior.
Approach: Examine Cross-reference the list of users with the total number of photos likes

```
153
154 •  SELECT users.id, username, COUNT(users.id) AS total_num_of_likes
155    FROM users
156    JOIN likes ON users.id = likes.user_id GROUP BY users.id
157    HAVING total_num_of_likes = (select count(*) from photos);
158
```

QUERY

Result Grid | Filter Rows:

id	username	total_num_of_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

OUTPUT

RESULT

This project, improved my SQL skills and demonstrated the importance of data-driven decisions in enhancing user experience and driving business growth on Instagram.

Through this project, I achieved several key milestones: identifying the five oldest users on Instagram, detecting users who have never posted a single photo, determining the contest winner with the most liked photo, and uncovering the top five most commonly used hashtags on the platform. These insights provided actionable information for the marketing team to target loyal users, re-engage inactive users, and optimize promotional strategies. Additionally, by calculating the average number of posts per user and identifying potential bot accounts, I offered critical insights into overall user engagement and the integrity of the user base

The insights derived from this project will support me in understanding strategic planning, ultimately contributing to success and growth.

Thank you

SAMIYA ALAM

samiyaalam1710@gmail.com