Instagram User Analytics

Overview

In this project, we leveraged SQL and MySQL Workbench to analyze user interactions and engagement on Instagram. The goal was to provide valuable insights to various teams within the organization, including marketing and product management, to drive informed decision-making and improve user experience.

Approach

First I downloaded and installed my SQL server and my SQL workbench on my system and installed it. Then I Created the necessary database ig_clone and all tables users, photos, comments, likes, follows, tags, photo_tags(you can see all the columns in ERD Diagram in next page.) and inserted all the values based on the provided Code.

Tools Used

1. MYSQL Server
2. MYSQL Workbench 8.0 CE
3. MYSQL Shell

Commands Used in this Project

- SELECT
- GROUP BY
- ORDER BY
- COUNT()
- CREATE
- INSERT INTO
- LIMIT
- SUM()
- JOINS(LEFT JOIN, RIGHT JOIN, INNER JOIN etc.)

Key Tasks and Objectives

1. Marketing Analysis:

   Loyal User Reward: Identify the five oldest users on Instagram to reward them.

   Inactive User Engagement: Find users who have never posted a single photo to encourage their activity.

   Contest Winner Declaration: Determine the user with the most likes on a single photo to declare the contest winner.

   Hashtag Research: Identify the top five most commonly used hashtags to guide promotional strategies.

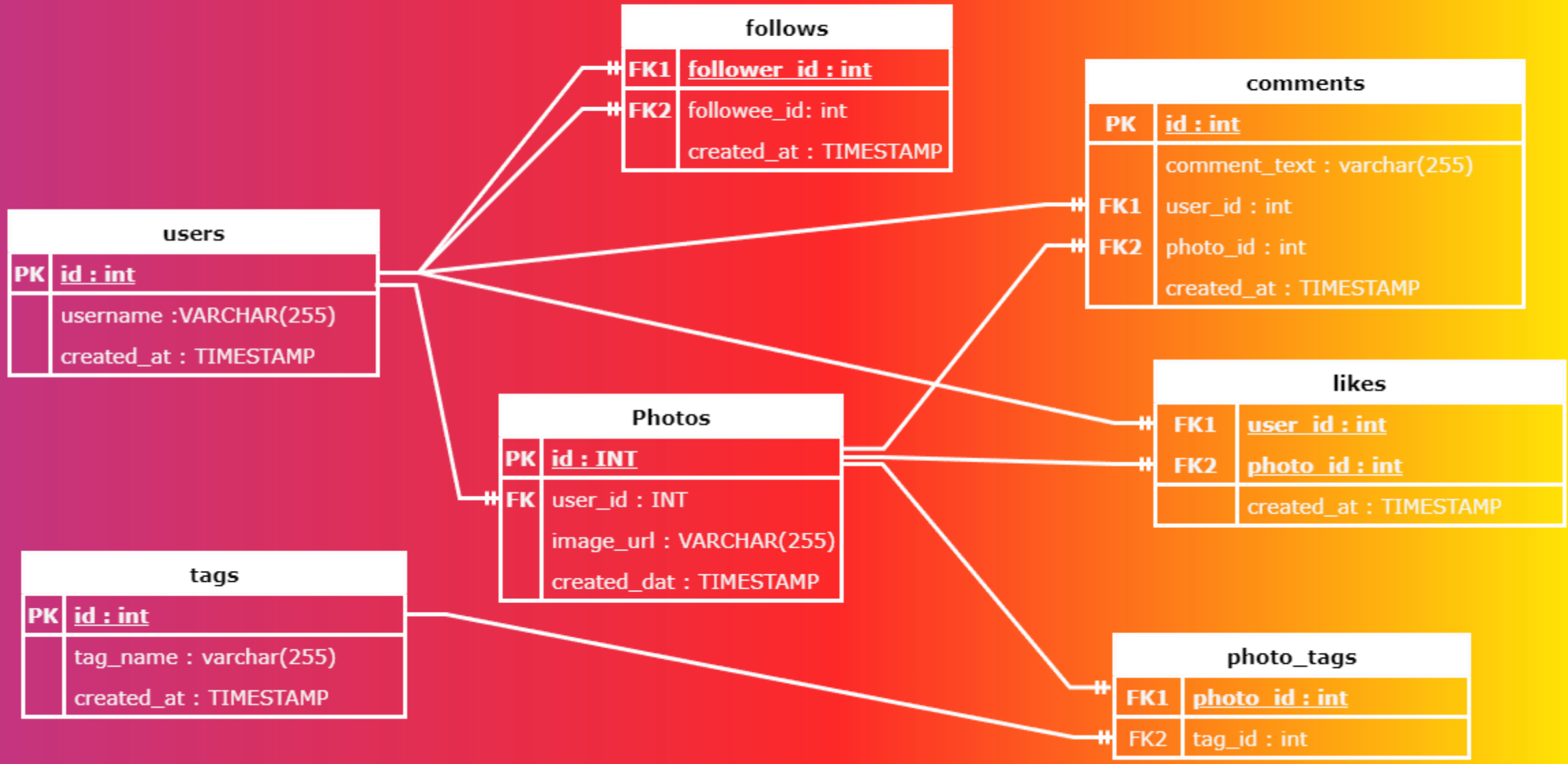
   Ad Campaign Launch: Determine the best day of the week for ad launches based on user registration data.

2. Investor Metrics:

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

   Bots & Fake Accounts: Identify potential bot accounts by flagging users who have liked every single photo on the platform.

Database : ig_clone



Loyal User Reward

Question - The marketing team wants to reward the most loyal users, i.e., those who have been using the platform for the longest time. Identify the five oldest users on Instagram from the provided database.

Input -

```
select * from users
order by created_at
limit 5;
```

Output -

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

Inactive User Engagement

Question - The team wants to encourage inactive users to start posting by sending them promotional emails. Identify users who have never posted a single photo on Instagram.

Input-

```
/*Inactive user engagement*/
SELECT users.id, username, created_at FROM photos
right join users
on photos.user_id=users.id
where image_url is null;
```

Output-

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
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.Bernhard	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

Contest winner Declaration

Question- The team has organized a contest where the user with the most likes on a single photo wins. Determine the winner of the contest and provide their details to the team.

Input-

```sql
/*Contest Winner Declaration*/
select username, photo_id, count(likes.user_id) as Total_likes from likes
inner join photos on photos.id=likes.photo_id
inner join users on users.id=photos.user_id
group by photo_id order by count(user_id) desc
Limit 1;
```

Output-

username	photo_id	Total_likes
Zack_Kemmer93	145	48

Hashtag Research

Question- A partner brand wants to know the most popular hashtags to use in their posts to reach the most people. Identify and suggest the top five most commonly used hashtags on the platform.

Input-

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

Output-

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

Ad Campaign Launch

Question -The team wants to know the best day of the week to launch ads. Determine the day of the week when most users register on Instagram. Provide insights on when to schedule an ad campaign.

Input -

```
/* Ad Campaign Launch*/
select dayname(created_at) as day_of_the_week,
count(dayname(created_at)) as user_registered from users
group by dayname(created_at)
order by count(dayname(created_at)) desc;
```

Output -

day_of_the_week	user_registered
Thursday	16
Sunday	16
Friday	15
Tuesday	14
Monday	14
Wednesday	13
Saturday	12

User Engagement

Question- Investors want to know if users are still active and posting on Instagram or if they are making fewer posts. 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.

Input -

```
/* User Engagement */
select (count(image_url)/count(distinct(user_id))) as avg_post_per_user,
(count(image_url)/count(users.id)) as total_post_by_total_user  from photos
inner join users on users.id = photos.user_id;
```

Output -

avg_post_per_users	total_post_by_total_user
3.4730	1.0000

Bots & Fake Accounts

Question - Investors want to know if the platform is crowded with fake and dummy accounts. Identify users (potential bots) who have liked every single photo on the site, as this is not typically possible for a normal user.

Input -

```sql
/* Bots & Fake Accounts */
select user_id,username, count(user_id) as Total_post_liked from likes
join users on users.id = likes.user_id
group by user_id
having count(user_id) = 257;
```

Output -

user_id	username	Total_post_liked
5	Aniya_Hackett	257
14	Jadyn81	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

Results

1. Darby_Herzog, Emilio_Bernier52, Elenor88, Nicole71 and Jordyn_Jacobson2 are five of the oldest users of the platform.
2. users who have never posted a single photo on Instagram are Aniya_Hackett, Kasandra_Homenick, Jaclyn81, Rocio33, Makwell.Halvorson, Ticya.Trantow, Pearl7, Ollie_Ledner37, Mckenna17, David.Osinski47, Morgan.Kassulke, Linnea59, Duane60, Julien_Schmidt, Mike.Auer39, Franco_Keebler64, Nia_Haag, Hulda.Macejkovic, Leslie67, Janelle. Nikolaus81, Darby Herzog, Esther. Zulauf61, Bartholome.Bernhard, Jessyca_West, Esmeralda.Mraz57 and Bethany20.
3. Zeck_kemmer93 the user with the most likes on a single photo with 48 likes on a single photo.
4. the top five most commonly used hashtags on the platform are smile which has been used 59 times, beach, party, fun and concert are used 42, 39, 38, 38 and 24 times respectively.

5. Thursday and Sunday can be good to schedule a campaign as most users register on these days on the platform.

6. The average post per user is 3, whereas total post by total user is 1, which shows some people who are registered do not very often.

7. Users with likes on every post on Instagram are Aniya_Hackett, Jaclyn81, Rocio33, Maxwell.Halvorson, Ollie_Ledner37, Mckenna17, Duane60, Julien_Schmidt, Mike. Auer39, Nia_Haag, Leslie67, Janelle.Nikolaus81 and Bethany20.

Insights

Knowledge Gained while working on this project

- Through this Project I learnt many important concepts of solving SQL queries.
- It sharpen my SQL concepts specifically joins, I learnt how to use joins on more than two tables and it gave me a clarity about how to use group by command correctly.