Instagram User Analytics

~Vineet Patel

Project Description: In this project, I as a data analyst within the Instagram product team, I am tasked with analysing user interactions and engagement with the Instagram app using ig_clone Database to provide valuable insights that can help the business grow.

The database consists of total 7 tables,.

1. users
2. follows
3. photos
4. tags
5. photo_tags
6. likes
7. comments

I'll be extracting meaningful insights from the above database to help the product manager and the rest of the team make informed decisions about the future direction of the Instagram app.

Approach: My approach towards the project is first I created ig_clone database which was provided to me with the project. After that I wrote all the queries to find the solution of the question which was asked.

Tach stack Used : I used SQL(Version 8.0.34) which have downloaded from link provided and then after I downloaded the SQL

Workbench(Version 8.0.34) I chose MySQL for the project because it is fast and easy to use.

Insights :

A) Marketing Analysis:

1. Loyal User Reward: The marketing team wants to reward the most loyal users, i.e., those who have been using the platform from the longest time.

Task: Identify the five oldest users on Instagram from the provided dataset.

Query:

```
1    # the five oldest users on Instagram
2 •  USE ig_clone;
3 •  SELECT *
4    FROM users
5    ORDER BY created_at
6    LIMIT 5;
```

Result:

2.

	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: The team wants to encourage inactive users to start posting by sending them promotional emails.

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

Query:

```
10      # users who have never posted a single photo on Instagram.
11  •   USE ig_clone;
12  •   SELECT * FROM users;
13  •   SELECT * FROM photos;
14  •   SELECT  username ,
15              user_id, image_url
16      FROM users
17      LEFT JOIN photos
18          ON users.id = photos.user_id
19      Where photos.id IS NULL;
```

Result:

username	user_id	image_url
Aniya_Hackett	NULL	NULL
Bartholome.Bernhard	NULL	NULL
Bethany20	NULL	NULL
Darby_Herzog	NULL	NULL
David.Osinski47	NULL	NULL
Duane60	NULL	NULL
Esmeralda.Mraz57	NULL	NULL
Esther.Zulauf61	NULL	NULL
Franco_Keebler64	NULL	NULL
Hulda.Macejkovic	NULL	NULL
Jaclyn81	NULL	NULL
Janelle.Nikolaus81	NULL	NULL
Jessyca_West	NULL	NULL

username	user_id	image_url
Julien_Schmidt	NULL	NULL
Kasandra_Homenick	NULL	NULL
Leslie67	NULL	NULL
Linnea59	NULL	NULL
Maxwell.Halvorson	NULL	NULL
Mckenna17	NULL	NULL
Mike.Auer39	NULL	NULL
Morgan.Kassulke	NULL	NULL
Nia_Haag	NULL	NULL
Ollie_Ledner37	NULL	NULL
Pearl7	NULL	NULL
Rocio33	NULL	NULL
Tierra.Trantow	NULL	NULL

There are 26 Users who have never posted a single photo on Instagram.

4.

Contest Winner Declaration: The team has organized a contest where the user with the most likes on a single photo wins.

Task: Determine the winner of the contest and provide their details to the team.

Query:

```
# user with the most likes on a single photo
USE ig_clone;
SELECT * FROM photos;
SELECT * FROM likes;
SELECT
    users.id AS user_id,
    username,
    photos.id AS photo_id,
    photos.image_url,
    count(*) AS total_likes
FROM photos
    INNER JOIN likes
        ON likes.photo_id = photos.id
    INNER JOIN users
        ON photos.user_id = users.id
GROUP BY photos.id
ORDER BY total_likes DESC
LIMIT 1;
```

Result:

	user_id	username	photo_id	image_url	total_likes
▶	52	Zack_Kemmer93	145	https://jarret.name	48

5.

Hashtag Research: A partner brand wants to know the most popular hashtags to use in their posts to reach the most people.

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

Query:

```
50    # top five most commonly used hashtags on the platform.
51 •  USE ig_clone;
52 •  SELECT * FROM tags;
53 •  SELECT * FROM photo_tags;
54 •  SELECT tags.tag_name,
55        COUNT(*) AS tag_count
56    FROM photo_tags
57        JOIN tags
58            ON photo_tags.tag_id = tags.id
59    GROUP BY tags.id
60    ORDER BY tag_count DESC
61    LIMIT 7;
```

Result:

	tag_name	tag_count
▶	smile	59
	beach	42
	party	39
	fun	38
	food	24
	lol	24
	concert	24

24

6.

Top 4 most commonly used hashtags are (smile, beach, party, fun)

And the are three hashtags at 5th place (food, lol, concert)

Ad Campaign Launch: The team wants to know the best day of the week to launch ads.

Task: Determine the day of the week when most users register on Instagram. Provide Insights on when to schedule an ad campaign.

Query:

```
68    # the day of the week when most users register on Instagram
69 •  USE ig_clone;
70 •  SELECT
71        DAYNAME(created_at) AS `day`,
72        COUNT(*) AS day_count
73    FROM users
74    GROUP BY `day`
75    ORDER BY day_count DESC
76    LIMIT 2;
```

Result:

	day	day_count
▶	Thursday	16
	Sunday	16

7.

Thursday and Sunday has the most numbers of users register on Instagram.

B) Investor Metrics:

1. **User Engagement:** Investors want to know if users are still active and posting on Instagram or if the are making fewer posts.

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.

Query:

```
82    #average number of posts per user on Instagram
83 •  USE ig_clone;
84 •  SELECT (SELECT COUNT(*) FROM photos)
85          / (SELECT COUNT(*) FROM users)
86    AS avg_num_of_posts_per_user;
```

Result:

	avg_num_of_posts_per_user
▶	2.5700

2. **Bots & Fake Accounts:** Investors want to know if the platform is crowded with fake and dummy accounts.

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

Query:

```
94    # Identify users (potential bots) who have liked every single photo on the site
95    USE ig_clone;
96    SELECT
97            username,
98            user_id,
99            COUNT(*) AS total_likes
100   FROM users
101   INNER JOIN likes
102   ON users.id = likes.user_id
103   GROUP BY likes.user_id
104   HAVING total_likes = (SELECT COUNT(*) FROM photos);
```

Result:

username	user_id	total_likes
Aniya_Hackett	5	257
Jadyn81	14	257
Rocio33	21	257
Maxwell.Halvorson	24	257
Ollie_Ledner37	36	257
Mckenna17	41	257
Duane60	54	257
Julien_Schmidt	57	257
Mike.Auer39	66	257
Nia_Haag	71	257
Leslie67	75	257
Janelle.Nikolaus81	76	257
Bethany20	91	257

Overall Result: While doing this project I got to learn about using SQL and MySQL Workbench and I believe that it will help me a lot in future. I got to know how to solve real world problems using MySQL Query. I have provided the solutions of every questions asked and I believe they are correct.