INSTAGRAM USER ANALYTICS

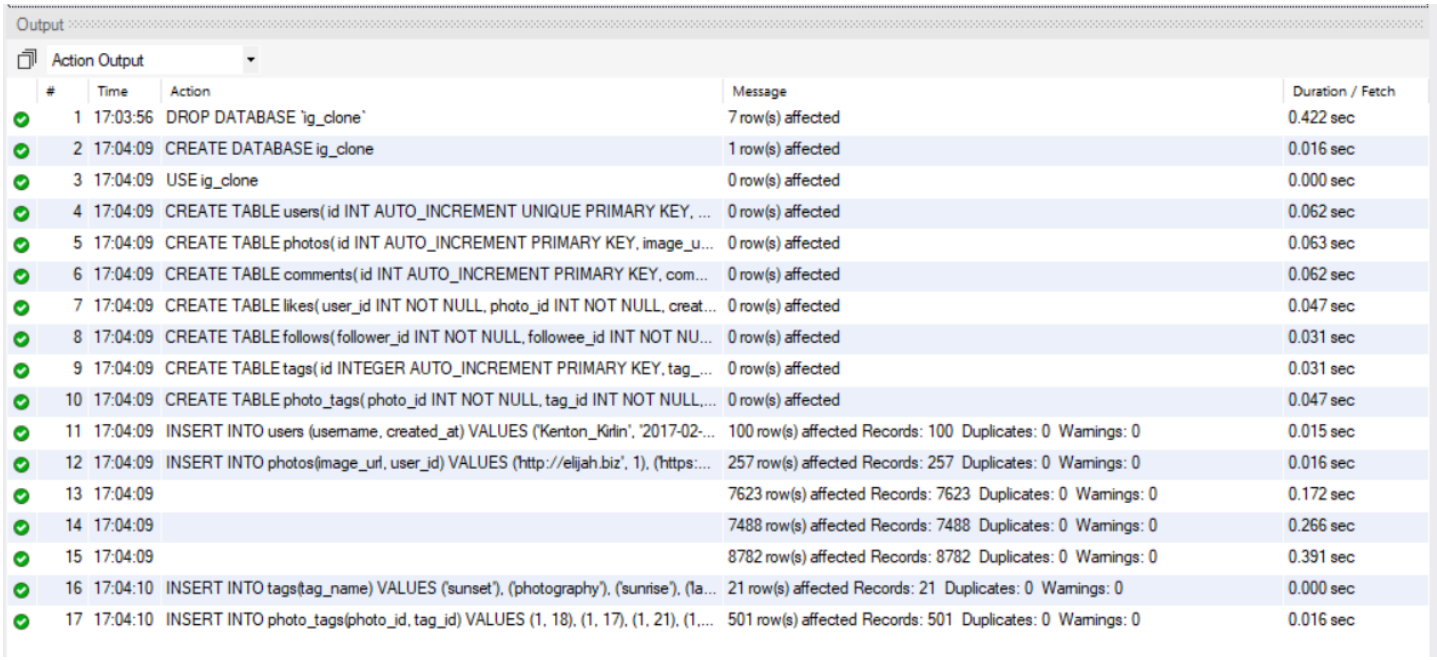
Project description

An Instagram user analytics project typically involves gathering and analyzing data related to user interactions, engagement, and demographics on the platform.

The goal of such a project is to gain insights into the behavior of an Instagram account's audience, understand what content resonates the most with them, identify trends, and optimize future content strategies

Approach

For this project, I used the dataset provided by the Trainity team to create the required tables and populate them accordingly in MySQL. The various queries I used to load the data and gain insights are mentioned in the results section



Output

Action Output

#	Time	Action	Message	Duration / Fetch
1	17:03:56	DROP DATABASE `ig_clone`	7 row(s) affected	0.422 sec
2	17:04:09	CREATE DATABASE ig_clone	1 row(s) affected	0.016 sec
3	17:04:09	USE ig_clone	0 row(s) affected	0.000 sec
4	17:04:09	CREATE TABLE users(id INT AUTO_INCREMENT UNIQUE PRIMARY KEY, ...	0 row(s) affected	0.062 sec
5	17:04:09	CREATE TABLE photos(id INT AUTO_INCREMENT PRIMARY KEY, image_u...	0 row(s) affected	0.063 sec
6	17:04:09	CREATE TABLE comments(id INT AUTO_INCREMENT PRIMARY KEY, com...	0 row(s) affected	0.062 sec
7	17:04:09	CREATE TABLE likes(user_id INT NOT NULL, photo_id INT NOT NULL, creat...	0 row(s) affected	0.047 sec
8	17:04:09	CREATE TABLE follows(follower_id INT NOT NULL, followee_id INT NOT NU...	0 row(s) affected	0.031 sec
9	17:04:09	CREATE TABLE tags(id INTEGER AUTO_INCREMENT PRIMARY KEY, tag_...	0 row(s) affected	0.031 sec
10	17:04:09	CREATE TABLE photo_tags(photo_id INT NOT NULL, tag_id INT NOT NULL,...	0 row(s) affected	0.047 sec
11	17:04:09	INSERT INTO users (username, created_at) VALUES ('Kenton_Kirlin', '2017-02-...	100 row(s) affected Records: 100 Duplicates: 0 Warnings: 0	0.015 sec
12	17:04:09	INSERT INTO photos(image_url, user_id) VALUES ('http://elijah.biz', 1), ('https:...	257 row(s) affected Records: 257 Duplicates: 0 Warnings: 0	0.016 sec
13	17:04:09		7623 row(s) affected Records: 7623 Duplicates: 0 Warnings: 0	0.172 sec
14	17:04:09		7488 row(s) affected Records: 7488 Duplicates: 0 Warnings: 0	0.266 sec
15	17:04:09		8782 row(s) affected Records: 8782 Duplicates: 0 Warnings: 0	0.391 sec
16	17:04:10	INSERT INTO tags(tag_name) VALUES ('sunset'), ('photography'), ('sunrise'), ('la...	21 row(s) affected Records: 21 Duplicates: 0 Warnings: 0	0.000 sec
17	17:04:10	INSERT INTO photo_tags(photo_id, tag_id) VALUES (1, 18), (1, 17), (1, 21), (1,...	501 row(s) affected Records: 501 Duplicates: 0 Warnings: 0	0.016 sec

Tech Stack used

For this project, I have chosen MySQL for my database management software as it is the most widely used open-source relational database management system known for its stability and reliability. I have also used MySQL for a number of college projects hence it was my first choice

MySQL is optimized for performance, providing fast read and write operations, which is crucial for real-time analytics applications

Insights

Loyal User Reward : The marketing team wants to reward the most loyal users, i.e., those who have been using the platform for the longest time. Your Task: Identify the five oldest users on Instagram from the provided database.

```
89    # Identify the five oldest users on Instagram from the provided database.
90    SELECT * FROM users ORDER BY created_at LIMIT 5;
91
```

Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell

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

Therefore we found out that the five oldest users are Darby Herzog, Emilio Bernier, Elenor, Nicole and Jordyn Jacobson

Inactive User Engagement : The team wants to encourage inactive users to start posting by sending them promotional emails. Your Task: Identify users who have never posted a single photo on Instagram.

```
94    #Identify users who have never posted a single photo on Instagram.
95    SELECT username
96    FROM users
97    LEFT JOIN photos
98    ON users.id = photos.user_id
99    WHERE photos.id IS NULL;
100
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

username
Aniya_Hackett
Kasandra_Homenick
Jaclyn81
Rocio33
Maxwell.Halvorson
Tierra.Trantow
Pearl7
Ollie_Ledner37
Mckenna17
David.Osinski47

There the above users have never posted a single photo on instagram

Contest Winner Declaration : The team has organized a contest where the user with the most likes on a single photo wins. Your Task: Determine the winner of the contest and provide their details to the team.

```
103     # Determine the winner of the contest and provide their details to the team.
104 •   SELECT username, photos.id, photos.image_url, count(likes.user_id) AS total
105     FROM photos
106     INNER JOIN likes
107     ON likes.photo_id = photos.id
108     INNER JOIN users
109     ON photos.user_id = users.id
110     GROUP BY photos.id
111     ORDER BY total DESC
112     LIMIT 1;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

username	id	image_url	total
Zack_Kemmer93	145	https://jarret.name	48

Therefore the winner of the contest is Zack Kemmer as he has the highest number of likes

Hashtag Research: A partner brand wants to know the most popular hashtags to use in their posts to reach the most people. Your Task: Identify and suggest the top five most commonly used hashtags on the platform.

```
114     # Identify and suggest the top five most commonly used hashtags on the platform.
115 •   SELECT tags.tag_name, COUNT(*) AS total
116     FROM photo_tags
117     JOIN tags
118     ON photo_tags.tag_id = tags.id
119     GROUP BY tags.id
120     ORDER BY total DESC
121     LIMIT 5;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

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

Therefore the top five most commonly used hashtags are smile, beach, party, fun and concert

Ad Campaign Launch: The team wants to know the best day of the week to launch ads. Your Task: Determine the day of the week when most users register on Instagram. Provide insights on when to schedule an ad campaign.

```
124    # Determine the day of the week when most users register on Instagram.
125    #Provide insights on when to schedule an ad campaign.
126    SELECT DAYNAME (created_at) AS day, count(*) as total
127    FROM users
128    GROUP BY day
129    ORDER BY total DESC
130    LIMIT 1;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

day	total
Thursday	16

Therefore the best day to launch ads is thursday

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

```
133    # Calculate the average number of posts per user on Instagram.
134    # Also, provide the total number of photos on Instagram divided by the total number of users.
135    SELECT (SELECT COUNT(*) FROM photos)/(SELECT COUNT(*) FROM users) AS avg;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

avg
2.5700

The average number of posts per user is 2.57

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

```
136     # Identify users (potential bots) who have liked every single photo on the site,
137     # as this is not typically possible for a normal user.
138 •   SELECT user_id, COUNT(*) as num_likes
139     FROM likes
140     GROUP BY user_id
141     HAVING num_likes= (SELECT COUNT(*) FROM photos);
142 •   SELECT u.username, COUNT(*) as num_likes
143     FROM users u
144     JOIN likes l ON u.id = l.user_id
145     GROUP BY u.id
146     HAVING num_likes= (SELECT COUNT(*) FROM photos);
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

username	num_likes
Aniya_Hackett	257
Jadyn81	257
Rocio33	257
Maxwell.Halvorson	257
Ollie_Ledner37	257
Mckenna17	257
Duane60	257
Julien_Schmidt	257
Mike.Auer39	257
Nia_Haag	257

Therefore the above users are potential bots

Results

Hence we were able to the MySQL database management software to fire multiple queries that gave us valuable insights on instagram user analytics