INSTAGRAM USER ANALYTICS REPORT

PROJECT DESCRIPTION

1. This project uses **SQL** to analyze user data from a realistic **Instagram clone database**, providing insights into user engagement, loyalty, content patterns, and potential fraud.
2. The analysis supports business decisions in marketing, product development, and investor relations.

APPROACH

1. **Database Setup:** Created and validated the MySQL database using the provided schema, loaded sample data, and performed initial quality checks to ensure data integrity.
2. **Stakeholder Alignment:** Gathered requirements from marketing, product, and investor teams to define key questions and align analysis goals with business priorities.
3. **Query Development:** Designed precise, modular SQL queries for each task, focusing on clarity, efficiency, and performance. Used subqueries, joins, and aggregates as needed.
4. **Exploratory Data Analysis:** Conducted exploratory analysis to identify trends, outliers, and user segments. Examined distributions, time series, and engagement patterns.
5. **Insight Generation:** Turned SQL results into actionable insights—highlighting loyal users, inactive accounts, content winners, popular hashtags, and potential bots. Provided clear recommendations for each finding.
6. **Documentation:** Maintained a query library and documented assumptions, processes, and limitations for reproducibility and team knowledge sharing.

TECH-STACK USED

1. **Database:** MySQL Ver 8.0.42
2. **Tool:** MySQL Workbench
3. **Why:** MySQL is widely used for relational data analysis, and Workbench offers a robust GUI for querying and result visualization.

SQL QUERIES AND INSIGHTS

```
USE ig_clone;
```

A) Marketing Analysis

1. LOYAL USER REWARD: Identify the five oldest users

QUERY:

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

RESULT:

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

INSIGHT:

Reward these users as Instagram's earliest adopters. Their loyalty can be leveraged in testimonials or referral programs.

2. INACTIVE USER ENGAGEMENT: Users who never posted a photo

QUERY:

```
SELECT id, username
FROM users
WHERE id NOT IN (SELECT DISTINCT user_id FROM photos);
```

RESULT:

id	username
5	Aniya_Hackett
7	Kasandra_Homenick
14	Jaclyn81
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

INSIGHT:

Target these users with onboarding campaigns or incentives to encourage their first post.

3. CONTEST WINNER DECLARATION: User with most likes on a single photo

QUERY:

```
SELECT users.id, users.username, photos.id AS photo_id,
COUNT(likes.user_id) AS like_count
FROM likes
JOIN photos ON likes.photo_id = photos.id
JOIN users ON photos.user_id = users.id
GROUP BY photo_id
ORDER BY like_count DESC
LIMIT 1;
```

RESULT:

id	username	photo_id	like_count
52	Zack_Kemmer93	145	48

INSIGHT:

Promote the winner's content and consider similar contests to boost engagement.

4. HASHTAG RESEARCH: Top five most used hashtags

QUERY:

```
SELECT tags.tag_name, COUNT(photo_tags.tag_id) AS usage_count
FROM tags
JOIN photo_tags ON tags.id = photo_tags.tag_id
GROUP BY tags.tag_name
ORDER BY usage_count DESC
LIMIT 5;
```

RESULT:

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

INSIGHT:

Suggest these hashtags to the partner brand for maximum reach.

5. AD CAMPAIGN LAUNCH: Best day for user registrations

QUERY:

```
SELECT DAYNAME(created_at) AS day_of_week, COUNT(*) AS signups
FROM users
GROUP BY day_of_week
ORDER BY signups DESC
LIMIT 1;
```

RESULT:

day_of_week	signups
Thursday	16

INSIGHT:

Schedule major ad campaigns on this day for higher visibility among new users.

B) Investor Metrics:

1. USER ENGAGEMENT: Average posts per user & Total photos divided by total users

QUERY (Average posts per user):

```
SELECT COUNT(photos.id) / COUNT(DISTINCT users.id) AS avg_posts_per_user
FROM users
LEFT JOIN photos ON users.id = photos.user_id;
```

RESULT:

avg_posts_per_user
2.5700

INSIGHT:

This metric shows user activity health—higher averages indicate more engaged users.

QUERY (Total photos divided by total users):

```
SELECT COUNT(*) AS total_photos,
       COUNT(DISTINCT user_id) AS total_users_with_posts,
        COUNT(*) / COUNT(DISTINCT user_id) AS photos_per_user
FROM photos;
```

RESULT:

total_photos	total_users_with_posts	photos_per_user
257	74	3.4730

INSIGHT:

Use this ratio to track content generation trends over time.

2. BOTS & FAKE ACCOUNTS: Users who liked every photo

QUERY:

```
SELECT user_id, COUNT(*) AS likes_given
FROM likes
GROUP BY user_id
HAVING likes_given = (SELECT COUNT(*) FROM photos);
```

RESULT:

user_id	likes_given
5	257
14	257
21	257
24	257
36	257
41	257
54	257
57	257
66	257
71	257
75	257
76	257
91	257

INSIGHT:

These accounts are likely bots—consider implementing CAPTCHA or additional verification.

RESULTS

1. **Actionable Insights:** Provided data-driven recommendations for marketing, product, and development teams.
2. **SQL Expertise:** Demonstrated advanced querying skills for complex analytics.
3. **Business Impact:** Empowered decision-makers with metrics critical for growth and fraud prevention.