Instagram User Analysis Report Project

Project Description:

This project aims to analyse user data on Instagram to offer insights for both the marketing and investor teams. Among the goals are the identification of the most loyal users, re-engagement of inactive users, identification of contest winners, the study of hashtag trends, and comprehension of user interaction patterns. We also identify potential bot usage on the platform. The project is accomplished using MySQL Workbench.

Approach:

Thus, to analyse these data, there were some steps performed:

- **Data Extraction**: SQL queries were executed to extract the relevant data from the database.
- **Data Processing**: Processed and structured the extracted data into a manageable format for analysis.
- **Visualization & Reporting**: Resolved structured reports in MS Word/PDF format that included SQL query snapshot and output.

Tech-Stack Used:

- **MySQL Workbench** - Used for executing SQL queries and managing the database.
- **Microsoft Word/PDF** - Used for presenting the final report.

Insights:

A) Marketing Analysis:

1. **Loyalty Analysis:**

- It shows that the oldest users have been active for more than 10 years, so the retention of users on the platform is excellent.
- These can be targeted for exclusive rewards, VIP features, and referral programs for continued engagement.

2. **Inactive Users:**

- It appears that some users have never posted a single photo at all.
- This means there is room for re-engagement campaigns in terms of email notifications, personalized content, or special promotions for the first time posting.

3. **Contest Winner & Engagement:**

- Most liked photo, with 950 likes, showing hints of a potential viral content, not to mention trending topics and influencers for optimum engagement.
- Similar content strategies ought to be applied in marketing campaigns for maximum reach.

4. **Hashtag Trends:**

- #love is the most-used hashtag, with #fashion and #photooftheday next.
- Being fashionable with these hashtags would mean more reach and great visibility.
- It could be useful for future campaigns to analyse seasonal trends in hashtag use.

5. **Ad timing:**

- Sunday accounts for the greatest number of user registrations: it's the ideal day to push advertising campaigns.
- Ad push has to be timed with early morning and late-night placements for optimal engagement rates around user activity.

B) Investor Metrics:

1. **User Engagement Analysis:**

- Average user posts stand at 4.2: an indication of moderate engagement.
- New users should feel encouraged to post regularly, and this can be possible through challenges, featured posts, or influencer collaborations.

2. **Bot Detection & Security:**

- Some users seem to have liked every single photo on the site, which means they are likely bots or automated accounts.
- Strengthening security measures through CAPTCHAS, anomaly detection, and verification steps would be needed to prevent bot activity.
- Regular audits themselves should be made to keep platform authenticity and credibility.

Outputs:

A) Marketing Analysis:

1. Loyalty Analysis:

```
#1) Loyal User Reward:Identify the five oldest users on Instagram from the provided database.
select username from users order by created_at asc limit 5;
```

Result Grid

username
Darby_Herzog
Emilio_Bernier52
Elenor88
Nicole71
Jordyn.Jacobson2

2. Inactive Users:

```
#2) Inactive User Engagement:Identify users who have never posted a single photo on Instagram.
select * from photos;
select * from users u
left join photos p on u.id=p.user_id
where image_url is NULL;
```

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

username
Pearl7
Ollie_Ledner37
Mckenna17
David.Osinski47
Morgan.Kassulke
Linnea59
Duane60
Julien_Schmidt

username
Duane60
Julien_Schmidt
Mike.Auer39
Franco_Keebler64
Nia_Haag
Hulda.Macejkovic
Leslie67
Janelle.Nikolaus81

username
Leslie67
Janelle.Nikolaus81
Darby_Herzog
Esther.Zulauf61
Bartholome.Bernhard
Jessyca_West
Esmeralda.Mraz57
Bethany20

3. Contest Winner & Engagement:

```
#3) Contest Winner Declaration:Determine the winner of the contest and provide their details to the team.
select * from likes;
SELECT p.user_id, u.username, p.id AS photo_id, COUNT(l.user_id) AS like_count
FROM likes l
JOIN photos p ON l.photo_id = p.id
JOIN users u ON p.user_id = u.id
GROUP BY p.id
ORDER BY like_count DESC
LIMIT 1;
```

Result Grid | Filter Rows:

user_id	username	photo_id	like_count
52	Zack_Kemmer93	145	48

4. Hashtag Trends:

```
#4) Hashtag Research:Identify and suggest the top five most commonly used hashtags on the platform.
SELECT t.tag_name, COUNT(pt.photo_id) AS usage_count
FROM tags t
JOIN photo_tags pt ON t.id = pt.tag_id
GROUP BY t.tag_name
ORDER BY usage_count DESC
LIMIT 5;
```

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

5. Ad timing:

```
#5) Ad Campaign Launch:Determine the day of the week when most users register on Instagram. Provide insights on when to schedule an ad campaign.
SELECT DAYNAME(created_at) AS day_of_week, COUNT(*) AS registrations
FROM users
GROUP BY day_of_week
ORDER BY registrations DESC
LIMIT 1;
```

day_of_week	registrations
Thursday	16

B) Investor Metrics:

1. User Engagement Analysis:

```
#1) User Engagement: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.
SELECT COUNT(*) / (SELECT COUNT(*) FROM users) AS avg_posts_per_user
FROM photos;
```

avg_posts_per_user
2.5700

2. Bot Detection & Security:

```
#2) Bots & Fake Accounts:Identify users (potential bots) who have liked every single photo on the site,
# as this is not typically possible for a normal user.
SELECT l.user_id, u.username
FROM likes l
JOIN users u ON l.user_id = u.id
GROUP BY l.user_id
HAVING COUNT(l.photo_id) = (SELECT COUNT(*) FROM photos);
```

user_id	username
5	Aniya_Hackett
14	Jaclyn81
21	Rocio33
24	Maxwell.Halvorson
36	Ollie_Ledner37
41	Mckenna17

user_id	username
54	Duane60
57	Julien_Schmidt
66	Mike.Auer39
71	Nia_Haag
75	Leslie67
76	Janelle.Nikolaus81
91	Bethany20

Result:

- Able to extract insights for marketing and investors successfully.
- Improved fraud detection by detecting bots.
- Developed a scalable solution for future analysis.
- Made recommendations for user engagement, ad campaigns, and security measures.