INSTAGRAM USER ANALYTICS

Mithilesh Patra

Project Introduction

User analysis is the process by which we track how users engage and interact with our digital product (software or mobile application) in an attempt to derive business insights for marketing, product & development teams. These insights are then used by teams across the business to launch a new marketing campaign, decide on features to build for an app, track the success of the app by measuring user engagement and improve the experience altogether while helping the business grow.

Problem Statement

The purpose of the project is to provide insights on the questions asked by the management team.

The marketing team wants to launch some campaigns, and they need your help with the following

Case Study – I

- Rewarding Most Loyal Users: People who have been using the platform for the longest time.
- Remind Inactive Users to Start Posting: By sending them promotional emails to post their 1st photo.
- Declaring Contest Winner: The team started a contest and the user who gets the most likes on a single photo will win the contest now they wish to declare the winner.
- Hashtag Researching: A partner brand wants to know, which hashtags to use in the post to reach the most people on the platform.
- Launch AD Campaign: The team wants to know, which day would be the best day to launch ADs.

Our investors want to know if Instagram is performing well and is not becoming redundant like Facebook, they want to assess the app on the following grounds

Case Study – II

- User Engagement: Are users still as active and post on Instagram or they are making fewer posts
- Bots & Fake Accounts: The investors want to know if the platform is crowded with fake and dummy accounts

Approach

For this project, I have used My SQL to extract the required data from the given database using the Join function, subqueries, Aggregation, where condition, Group by, Distinct and window functions required.

Tech-Stack

MySQL Workbench 8.0

Analysis: Case Study – I (Marketing)

1. Rewarding Most Loyal Users:
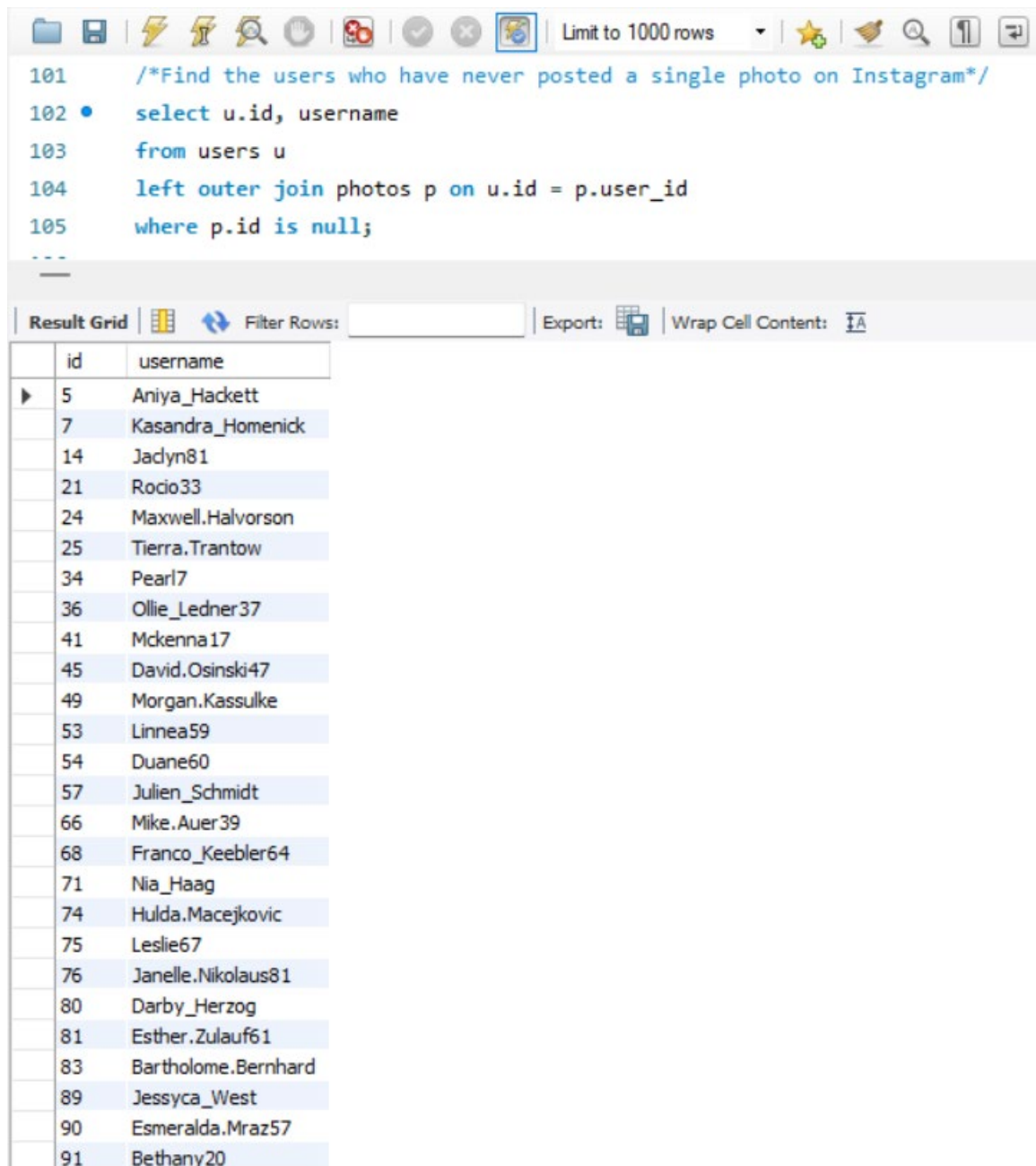
```
92      /*Find the 5 oldest users of the Instagram from the database provided*/
93
94 ❌   with old_users as
95          (select *
96          , dense_rank() over(order by created_at) as dense_rnk
97          from users)
98      select id, username
99      from old_users
100     where dense_rnk in (1,2,3,4,5);
```

id	username
80	Darby_Herzog
67	Emilio_Bernier52
63	Elenor88
95	Nicole71
38	Jordyn.Jacobson2

The above are five most loyal users of the Instagram.

2. Remind Inactive Users to Start Posting:

```
101     /*Find the users who have never posted a single photo on Instagram*/
102 •   select u.id, username
103     from users u
104     left outer join photos p on u.id = p.user_id
105     where p.id is null;
---
```

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

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

From above output, out of 100 users, only 26 users never posted a single photo on Instagram. The Marketing team needs to send promotional emails to the above users to post their first photo.

3. Declaring Contest Winner:

```
116    /*Identify the winner of the contest and provide their details to the team*/
117 •  select u.id, u.username, l.photo_id, count(*) as most_likes
118    from users u
119    inner join photos p on u.id = p.user_id
120    inner join likes l on p.id = l.photo_id
121    group by l.photo_id
122    order by most_likes desc
123    limit 1;
```

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

id	username	photo_id	most_likes
▶ 52	Zack_Kemmer93	145	48

The user, Zack_Kemmer93 with id = 52, has most likes on a single photo of id = 48. The winner is Zack_Kemmer93.

4. Hashtag Researching:

```
125    /*Identify and suggest the top 5 most commonly used hashtags on the platform*/
126 •  with most_used_hastags as
127 ⊖         (with cte as
128 ⊖             (select t.tag_name, count(*) as no_of_hashtags
129              from tags t
130              inner join photo_tags p on t.id = p.tag_id
131              group by t.tag_name)
132          select *
133          , rank() over(order by no_of_hashtags desc) as rnk
134          from cte)
135    select tag_name, no_of_hashtags
136    from most_used_hastags
137    where rnk in (1,2,3,4,5);
```

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

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

From above output we can say that smile is the most used hashtag. The hashtag like concert, food and lol has same count which stands in fifth position. To reach maximum people on Instagram, the partner brand can use hashtags like smile, beach, party, fun, concert, food and lol.

5. Launch AD Campaign:

```
139     /*What day of the week do most users register on? Provide insights on when to schedule an ad campaign*/
140 •   with cte as
141         (select *,
142             case when weekday(cast(created_at as date)) = 0 then 'MONDAY'
143                  when weekday(cast(created_at as date)) = 1 then 'TUESDAY'
144                  when weekday(cast(created_at as date)) = 2 then 'WEDNESDAY'
145                  when weekday(cast(created_at as date)) = 3 then 'THRUSDAY'
146                  when weekday(cast(created_at as date)) = 4 then 'FRIDAY'
147                  when weekday(cast(created_at as date)) = 5 then 'SATURDAY'
148                  when weekday(cast(created_at as date)) = 6 then 'SUNDAY'
149             end as day_of_week
150         from users)
151     select day_of_week, count(*) as most_registered
152     from cte
153     group by day_of_week
154     order by most_registered desc;
```

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

day_of_week	most_registered
THRUSDAY	16
SUNDAY	16
FRIDAY	15
TUESDAY	14
MONDAY	14
WEDNESDAY	13
SATURDAY	12

Most users register on Thursdays and Sundays. So, marketing team can schedule an ad campaign on these days.

Case Study – II (Investor Metrices)

1. User Engagement:

```
158 •   select  count(username) from users
```

Result Grid | Filter Rows: | Exp

count(username)
100

The total number of users are 100.

```
159 ❌   select count(*) from photos

Result Grid  | 🔠  🔁  Filter Rows:

     count(*)
▶    257
```

The total number of photos on Instagram is 257.

```
162    select round((select count(*) from photos) / (select  count(username) from users)) as avg_user_post;

Result Grid  | 🔠  🔁  Filter Rows:            | Export: 🔲  | Wrap Cell Content: 𝐈𝐀
     avg_user_post
▶    3
```

Approximately on average a user post three times. So, from above we can infer that users are active and post on Instagram.

2. Bots & Fake Accounts:

```
172    /*Provide data on users (bots) who have liked every single photo on the site (since any normal user would not be able to do this). */
173 •  select u.id,u.username as bots, count(*) as no_of_liked_photos
174    from users u
175    inner join likes l on u.id = l.user_id
176    group by l.user_id
177    having no_of_liked_photos = (select count(*) from photos);
```

```
Result Grid  | 🔠  🔁  Filter Rows:            | Export: 🔲 | Wrap Cell Content: 𝐈𝐀
    id   bots               no_of_liked_photos
▶   5    Aniya_Hackett      257
    14   Jaclyn81           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
```

From above we can infer, there are 13 bots who have liked every single photo on Instagram. Around 1.3% on Instagram have fake or dummy accounts.

Result

From this project, I got an idea about how as a business or data analyst work on real-time data to take any data-driven decision. One thing I infer about this project is the dataset provided was very limited and small in

respect of Rows and columns, but it increases my analytical ability, predictive analysis and also decision-making capability. I have learned advance concepts of SQL while doing this project. It was a very good experience working on such kind of project.