

Instagram User Analytics

Sakhi Patel

sakhipatel20@gmail.com

Project Description

This project aims to utilise SQL analysis to derive valuable insights from Instagram user data. By leveraging SQL queries, we seek to address various objectives outlined by the management team, including identifying loyal users, encouraging inactive users, determining contest winners, researching popular hashtags, optimizing ad campaign launch times, evaluating user engagement metrics, and detecting potential bot accounts. These insights will assist the product team in making informed decisions to enhance the Instagram app's performance and user experience.

Approach

- 1. Database Setup:** Create necessary database using provided commands and file.
- 2. Task Execution:** Address each objective systematically using SQL queries.
- 3. Analysis Documentation:** Document each SQL query and its output for transparency and clarity, utilizing MySQL Workbench for its robust features.

Tech-Stack Used: MySQL Workbench

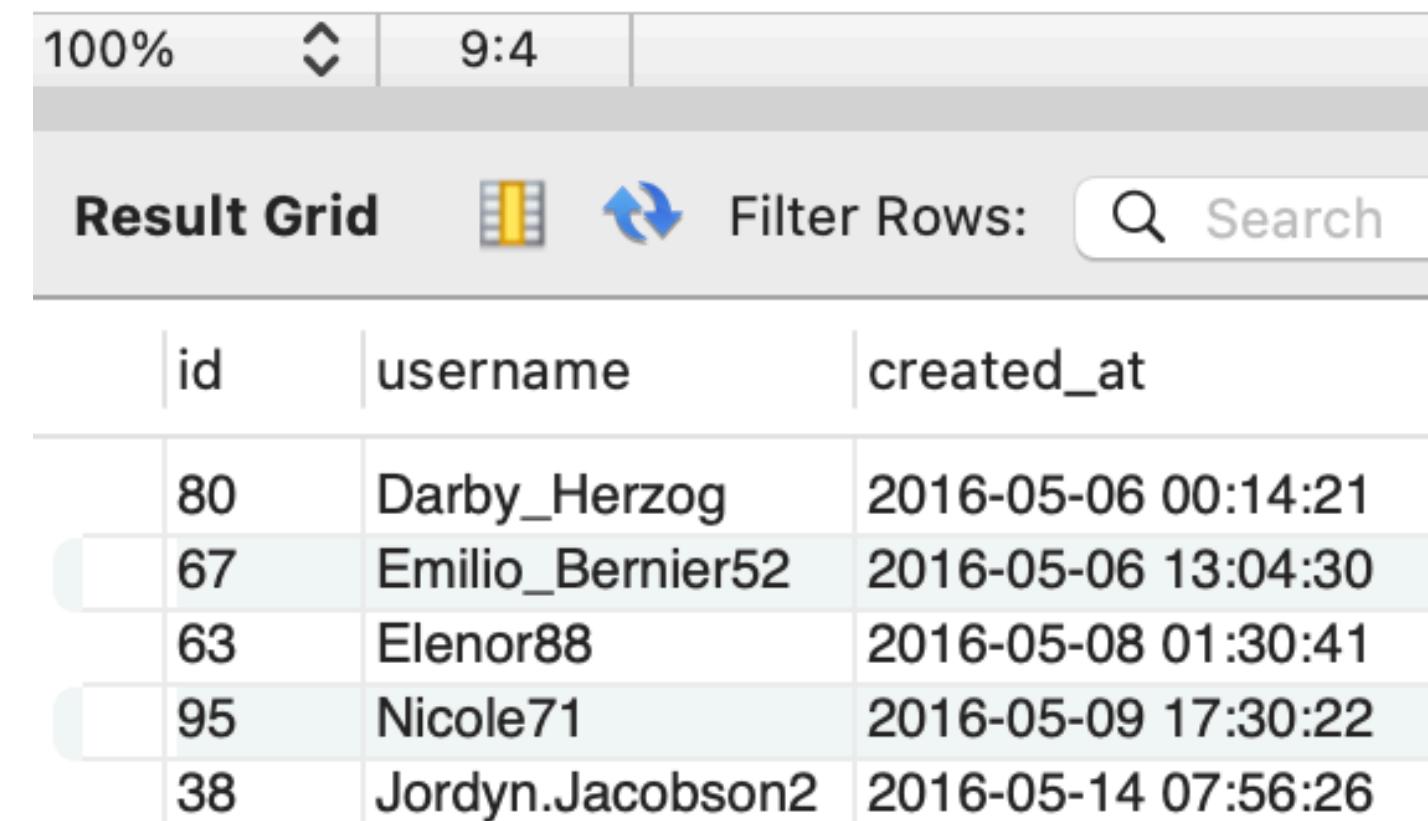
A) Marketing Analysis

1. Loyal User Reward
2. Inactive User Engagement
3. Contest Winner Declaration
4. Hashtag Research
5. Ad Campaign Launch

1. 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.

```
1 • select *
2   from users
3   order by created_at ASC
4   limit 5;
```



A screenshot of a database query results grid. The top bar shows zoom level (100%), orientation (9:4), and search/filter options. The results grid has columns for id, username, and created_at. The data shows five rows of user information, ordered by creation date.

	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

2. Inactive User Engagement

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

```
1 •   select username, users.id as user_id
2     from users
3   left join photos
4     on users.id = photos.user_id
5   where photos.id IS NULL
6   order by users.id;
```

2. Inactive User Engagement

Result Grid   Filter Rows

username	user_id
Aniya_Hackett	5
Kasandra_Homenick	7
Jaclyn81	14
Rocio33	21
Maxwell.Halvorson	24
Tierra.Trantow	25
Pearl7	34
Ollie_Ledner37	36
Mckenna17	41
David.Osinski47	45
Morgan.Kassulke	49
Linnea59	53
Duane60	54
Julien_Schmidt	57
Mike.Auer39	66
Franco_Keebler64	68
Nia_Haag	71
Hulda.Macejkovic	74
Leslie67	75
Janelle.Nikolaus81	76
Darby_Herzog	80
Esther.Zulauf61	81
Bartholome.Bernhard	83
Jessyca_West	89
Esmeralda.Mraz57	90
Bethany20	91

Out of the 100 users in our dataset, a notable portion of 26 users have never posted a single photo on Instagram.

3. Contest Winner Declaration

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

```
1 •  select users.id as user_id, users.username, photos.id as photo_id,  
2      photos.image_url, count(*) as total  
3      from photos  
4      inner join likes  
5      on likes.photo_id = photos.id  
6      inner join users  
7      on photos.user_id = users.id  
8      group by photos.id  
9      order by total DESC  
10     limit 1;
```

3. Contest Winner Declaration

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

Result Grid					
	user_id	username	photo_id	image_url	total
	52	Zack_Kemmer93	145	https://jarret.name	48

Zack_Kemmer93, user_id 52, wins the contest with photo_id 145, receiving the highest number of likes at 48.

4. Hashtag Research

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

```
1 • select tags.tag_name, count(*) as total_number_of_times_tag_used_individually
2   from tags
3   join photo_tags
4     on tags.id = photo_tags.tag_id
5   group by tags.tag_name
6   order by total_number_of_times_tag_used_individually DESC
7   limit 5;
```

4. Hashtag Research

Top 5 commonly used #Hashtags on Instagram

Result Grid   Filter Rows:

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

5. Ad Campaign Launch

The team wants to know the best day of the week to launch ads.

```
1 • select dayname(created_at) as day_of_week,  
2      count(*) as total_number_of_users_registered  
3      from users  
4      group by day_of_week  
5      order by total_number_of_users_registered DESC;
```

5. Ad Campaign Launch

The team wants to know the best day of the week to launch ads.

Result Grid   Filter Rows:  Search

	day_of_week	total_number_of_users_registered
	Thursday	16
	Sunday	16
	Friday	15
	Tuesday	14
	Monday	14
	Wednesday	13
	Saturday	12

Based on our analysis, a significant portion of user registrations occurred on Thursdays and Sundays, specifically on the 16th of the month. Therefore, launching ad campaigns on these two days would likely yield favorable results.

B) Investor Metrics

1. User Engagement
2. Bots & Fake Accounts

1. User Engagement

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

1. Times each user posts on Instagram:

```
1 • select user_id, count(*) as user_post_count
2   from photos
3   group by user_id
4   order by user_id;
```

1. User Engagement

Result Grid   Filter R

	user_id	user_post_count
1	1	5
2	2	4
3	3	4
4	4	3
6	6	5
8	8	4
9	9	4
10	10	3
11	11	5
12	12	4
13	13	5
15	15	4
16	16	4
17	17	3
18	18	1
19	19	2
20	20	1
22	22	1
23	23	12
26	26	5
27	27	1
28	28	4
29	29	8
30	30	2
31	31	1
32	32	4
33	33	5
35	35	2
37	37	1
38	38	2
39	39	1
40	40	1
42	42	3
43	43	5
44	44	4
46	46	4
47	47	5
48	48	1
50	50	3
51	51	5
52	52	5
55	55	1
56	56	1
58	58	8
59	59	10
60	60	2
61	61	1
62	62	2
63	63	4
64	64	5
65	65	5
67	67	3
69	69	1
70	70	1
72	72	5
73	73	1
77	77	6
78	78	5
79	79	1
82	82	2
84	84	2
85	85	2
86	86	9
87	87	4
88	88	11
92	92	3
93	93	2
94	94	1
95	95	2
96	96	3
97	97	2
98	98	1
99	99	3
100	100	2

1. User Engagement

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

2. Total number of photos divided by total number of users:

```
1 • select
2   (select count(*) from photos)/(select count(*) from users) as
3     total_photos_divide_total_users;
```

Result Grid			Filter Rows:	
total_photos_divide_total_users				
2.5700				

2. Bots & Fake Accounts

Investors want to know if the platform is crowded with fake and dummy accounts.

```
1 • select user_id, username, count(*) as total_likes_per_user
2   from users
3   inner join likes
4     on users.id = likes.user_id
5   group by likes.user_id, username
6   having total_likes_per_user = (select count(*) from photos);
```

2. Bots & Fake Accounts

Investors want to know if the platform is crowded with fake and dummy accounts.

Result Grid			
	user_id	username	total_likes_per_user
	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

These accounts have liked every available photo, indicating likely automated activity typical of bots.

Insights

Marketing Analysis:

- Identified the five earliest users for potential loyalty rewards.
- Queried for users with zero posts to target with reactivation campaigns.
- Determined the contest winner based on the highest post-like count.
- Uncovered the top five hashtags to guide partner brand recommendations.
- Analyzed user registration data to suggest the optimal day for launching ad campaigns.

Investor Metrics:

- Calculated the average post count per user to gauge user engagement.
- Established the photo-to-user ratio to understand content creation activity.
- Identified users with an abnormally high number of likes as potential bots for further investigation.

Results

- Utilized SQL analysis to tackle marketing and investor queries effectively.
- Delivered actionable insights for enhancing user engagement, content strategy, and ad campaigns.
- Improved comprehension of Instagram user behavior, aiding informed decision-making in the product team.



By: Sakhi Patel

Email ID: sakhipatel20@gmail.com