

Instagram User Analytics

Submission by Swet Parekh

Description:

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.

You are working with the product team of Instagram and the product manager has asked you to provide insights on the questions asked by the management team.

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

1. **Rewarding Most Loyal Users:** People who have been using the platform for the longest time.
Your Task: Find the 5 oldest users of the Instagram from the database provided

The screenshot shows a database query editor window titled "Query 1" with a toolbar and a "Limit to 1000 rows" option. The query is as follows:

```
1 • use ig_clone;  
2 • SELECT * FROM users  
3   ORDER BY created_at  
4   LIMIT 5;
```

Below the query editor is a "Result Grid" window showing the results of the query. The grid has columns for "id", "username", and "created_at". The results are as follows:

	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

At the bottom of the screenshot, there is a tab labeled "users 1" with a close button.

2. **Remind Inactive Users to Start Posting:** By sending them promotional emails to post their 1st photo.

Your Task: Find the users who have never posted a single photo on Instagram

```
5
6 • select username from users
7 left join photos on users.id=photos.user_id
8 where photos.id is null;
```

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

username
Aniya_Hackett
Kassandra_Homenick
Jadyn81
Rocio33
Maxwell_Halvorson
Tierra_Trantow
Pearl7
Ollie_Ledner37
Mckenna17
David_Osinski47
Morgan_Kassulke
Linnea59
Duane60
Julien_Schmidt
Mike_Auer39
Franco_Keebler64
Mia_Haan

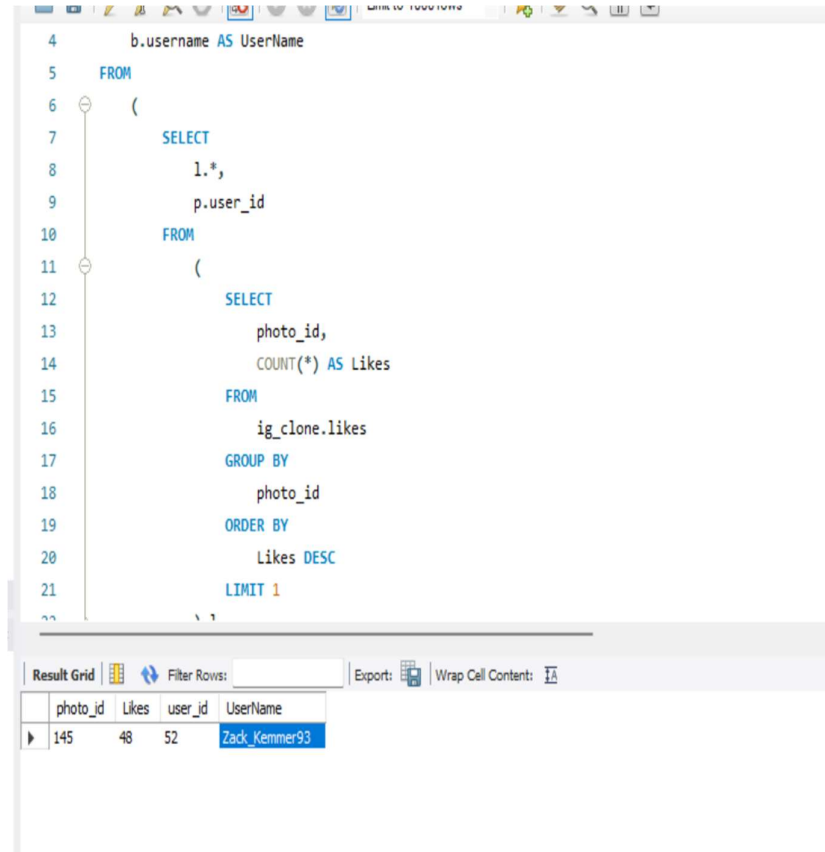
Result 2 x

Read Only

Input

#	Time	Action	Message	Duration / Fetch
2	01:06:53	SELECT * FROM users ORDER BY created_at LIMIT 5	5 row(s) returned	0.000 sec / 0.000 sec
3	01:27:20	select username from users left join photos on users.id=photos.user_id where photo...	26 row(s) returned	0.016 sec / 0.000 sec

3. **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.
- Your Task: Identify the winner of the contest and provide their details to the team



The screenshot shows a SQL query editor with a query that identifies the user with the most likes on a single photo. The query is as follows:

```
4      b.username AS UserName
5  FROM
6  (
7      SELECT
8          1.*,
9          p.user_id
10     FROM
11     (
12         SELECT
13             photo_id,
14             COUNT(*) AS Likes
15         FROM
16             ig_clone.likes
17         GROUP BY
18             photo_id
19         ORDER BY
20             Likes DESC
21         LIMIT 1
22     ) p
23 )
```

Below the query editor is a 'Result Grid' showing the output of the query. The grid has four columns: photo_id, Likes, user_id, and UserName. The first row of data shows photo_id 145, 48 Likes, user_id 52, and the username Zack_Kemmer93.

photo_id	Likes	user_id	UserName
145	48	52	Zack_Kemmer93

4. **Hashtag Researching:** A partner brand wants to know, which hashtags to use in the post to reach the most people on the platform.
Your Task: Identify and suggest the top 5 most commonly used hashtags on the platform

```
2 • SELECT
3     tag_name,
4     COUNT(*) AS hashtag_count
5 FROM
6     ig_clone.photo_tags
7 JOIN
8     ig_clone.tags ON ig_clone.photo_tags.tag_id = ig_clone.tags.id
9 GROUP BY
10    tag_name
11 ORDER BY
12    hashtag_count DESC
13 LIMIT 5;
14
15
```

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

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

Result 4 x

5. **Launch AD Campaign:** The team wants to know, which day would be the best day to launch ADs.
Your Task: What day of the week do most users register on? Provide insights on when to schedule an ad campaign

```
1 use ig_clone;
2 SELECT
3     DAYNAME(created_at) AS registration_day,
4     COUNT(*) AS user_count
5 FROM
6     ig_clone.users
7 GROUP BY
8     registration_day
9 ORDER BY
10    user_count desc
11
12 LIMIT 1;
13
14
```

Result Grid			Filter Rows:	Export:	Wrap Cell Cor
	registration_day	user_count			
▶	Thursday	16			

B) Investor Metrics: 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

- 1. **User Engagement:** Are users still as active and post on Instagram or they are making fewer posts
Your Task: Provide how many times does average user posts on Instagram. Also, provide the total number of photos on Instagram/total number of users

2 •
3
4
5
6
7
8
9
10

```
SELECT  
    COUNT(*) / (SELECT COUNT(*) FROM ig_clone.users) AS average_posts_per_user,  
    COUNT(*) AS total_photos,  
    (SELECT COUNT(*) FROM ig_clone.users) AS total_users  
FROM  
    ig_clone.photos;
```

Result Grid

Filter Rows:

Export:

Wrap Cell Content:

	average_posts_per_user	total_photos	total_users
▶	2.5700	257	100

2. **Bots & Fake Accounts:** The investors want to know if the platform is crowded with fake and dummy accounts
Your Task: 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).

```
2 • SELECT
3     user_id,
4     COUNT(*) AS total_likes
5 FROM
6     ig_clone.likes
7 GROUP BY
8     user_id
9 HAVING
10    total_likes = (SELECT COUNT(DISTINCT id) FROM ig_clone.photos);
11
```

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

	user_id	total_likes
▶	5	257
	14	257
	21	257
	24	257
	36	257
	41	257
	54	257
	57	257

Approach: For this project, I utilized a combination of SQL queries and data analysis techniques to gain insights into Instagram's user engagement and the presence of bots or fake accounts. I extracted relevant data from the database using SQL queries to answer specific questions posed by the investors. The data was then analyzed to provide meaningful insights regarding user activity, average posts per user, the total number of photos, and the identification of users who have liked every single photo on the site.

Tech-Stack Used:

I used SQL Workbench as the software to interact with the database and execute the SQL queries. SQL Workbench is a popular SQL client tool that provides a user-friendly interface for managing and querying databases. It allows for efficient data retrieval and analysis by providing a comprehensive set of features and functionalities. The queries were executed in SQL Workbench to extract the required data and derive insights from the database.

Insights: Through the project, I gained insights into various aspects of Instagram's performance. By calculating the average posts per user, I could assess user engagement and activity on the platform. Furthermore, by determining the total number of photos and users, I obtained a sense of the platform's overall size and reach. Additionally, identifying users who have liked every single photo on the site helped to highlight potential instances of fake or bot accounts, which is crucial for evaluating platform authenticity.

Result:

By utilizing SQL Workbench and executing the SQL queries, I successfully obtained insights into user engagement and the presence of bots or fake accounts on Instagram. The average posts per user metric provided an understanding of user activity, while identifying users who liked every photo helped to identify potential fake accounts. These insights enabled me to address investor concerns about the platform's performance and authenticity. The project's outcome was a comprehensive understanding of Instagram's user dynamics, which proved valuable for decision-making and assessing the platform's overall health.