

Instagram User Analytics

Description:

Imagine you're a data analyst working with the product team at Instagram. Your role involves analysing user interactions and engagement with the Instagram app to provide valuable insights that can help the business grow.

User analysis involves tracking how users engage with a digital product, such as a software application or a mobile app. The insights derived from this analysis can be used by various teams within the business. For example, the marketing team might use these insights to launch a new campaign, the product team might use them to decide on new features to build, and the development team might use them to improve the overall user experience. In this project, you'll be using SQL and MySQL Workbench as your tool to analyse Instagram user data and answer questions posed by the management team. Your insights will help the product manager and the rest of the team make informed decisions about the future direction of the Instagram app.

Remember, the goal of this project is to use your SQL skills to extract meaningful insights from the data. Your findings could potentially influence the future development of one of the world's most popular social media platforms.

SQL Tasks :

A) Marketing Analysis:

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.
Your Task: Identify the five oldest users on Instagram from the provided database.
2. **Inactive User Engagement:** The team wants to encourage inactive users to start posting by sending them promotional emails.
Your Task: Identify users who 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 win.

Your Task: Determine the winner of the contest and provide their details to the team.

4. **Hashtag Research:** A partner brand wants to know the most popular hashtags to use in their posts to reach the most people.

Your Task: Identify and suggest the top five most commonly used hashtags on the platform.

5. **Ad Campaign Launch:** The team wants to know the best day of the week to launch ads.

Your Task: Determine the day of the week when most users register on Instagram.
Provide insights on when to schedule an ad campaign.

B) Investor Metrics:

1. **User Engagement:** Investors want to know if users are still active and posting on Instagram or if they are making fewer posts.

Your Task: 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.

2. **Bots & Fake Accounts:** Investors want to know if the platform is crowded with fake and dummy accounts.

Your Task: Identify users (potential bots) who have liked every single photo on the site, as this is not typically possible for a normal user.

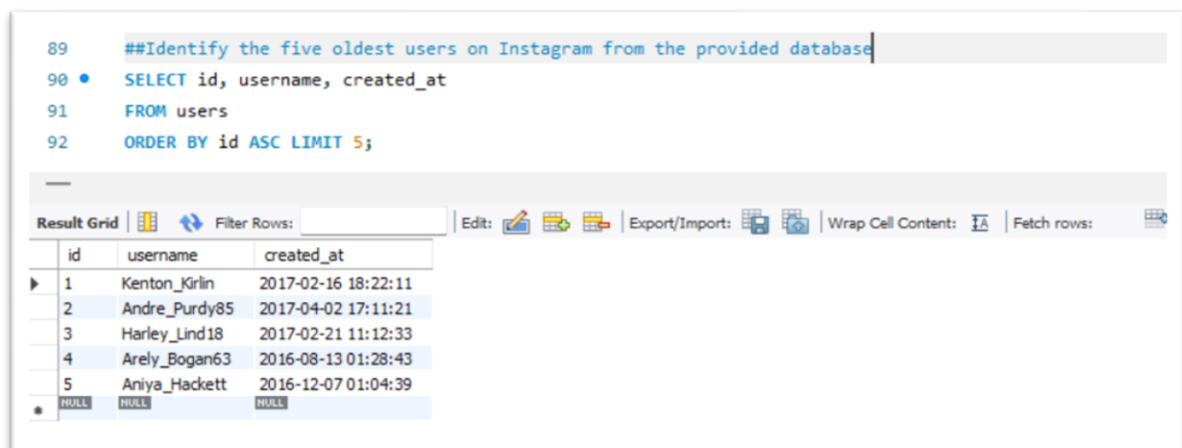
Solutions :

The project involved performing SQL queries on an Instagram user dataset. I am going to use MySQL workbench to analyse the data, user interactions and engagement. The approach included data analysis to understand the dataset structure, developing SQL queries to answer key business questions, data understanding to extract insights, and analysis of results and deriving useful insights.

I choose SQL and MySQL workbench, for their good database query performance and user-friendly interface. Additionally, SQL's built-in functions are useful to extract valuable insights.

- **Marketing Analysis:**

- 1) **Loyal User Reward:** Identified the five oldest users on Instagram, who have been active for the longest period.



The screenshot displays the MySQL Workbench interface. The SQL editor at the top contains the following query:

```
89  ##Identify the five oldest users on Instagram from the provided database
90  •  SELECT id, username, created_at
91     FROM users
92     ORDER BY id ASC LIMIT 5;
```

Below the editor, the 'Result Grid' tab is active, showing the results of the query. The results are presented in a table with three columns: 'id', 'username', and 'created_at'. The table contains five rows of data, with the first row highlighted in blue.

	id	username	created_at
▶	1	Kenton_Kirlin	2017-02-16 18:22:11
	2	Andre_Purdy85	2017-04-02 17:11:21
	3	Harley_Lind18	2017-02-21 11:12:33
	4	Arely_Bogan63	2016-08-13 01:28:43
	5	Aniya_Hackett	2016-12-07 01:04:39
*	NULL	NULL	NULL

- 2) Inactive User Engagement: Found users who have never posted a single photo, helping in targeted engagement strategies.

```
94  ##Identify users who have never posted a single photo on Instagram.
95
96  • SELECT
97      count(photos.id) as Photos_post,
98      photos.user_id,
99      users.username
100  from users
101  LEFT JOIN photos
102  ON users.id = photos.user_id
103  GROUP BY photos.user_id, users.username, users.id
104  HAVING photos_post = 0;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: [FA](#)

Photos_post	user_id	username
0	NULL	Aniya_Hackett
0	NULL	Kassandra_Homenick
0	NULL	Jadlyn81
0	NULL	Rocio33
0	NULL	Maxwell.Halvorson
0	NULL	Tierra.Trantow
0	NULL	Pearl7
0	NULL	Ollie_Ledner37
0	NULL	Mckenna17
0	NULL	David.Osinski47
0	NULL	Morgan.Kassulke
0	NULL	Linnea59
0	NULL	Duane60
0	NULL	Julien_Schmidt

- 3) Contest Winner Declaration: Determined the user with the highest likes on a single post.

```
107  ## Determine the winner of the contest and provide their details to the team:
108
109  • SELECT likes.user_id AS Total_likes, likes.photo_id, users.username FROM likes
110  JOIN users ON users.id = likes.user_id
111  JOIN photos ON likes.photo_id = photos.id
112  GROUP BY likes.user_id, likes.photo_id, users.username
113  ORDER BY Total_likes DESC LIMIT 1;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: [FA](#) | Fetch rows: [↕](#)

Total_likes	photo_id	username
100	255	Javonte83

- 4) Hashtag Research: Extracted the top five most commonly used hashtags to optimize brand marketing.

```
117 • SELECT COUNT(photo_tags.tag_id) AS common_Hashtag , tags.tag_name, tags.id FROM tags
118 LEFT JOIN photo_tags
119 ON tags.id = photo_tags.tag_id
120 GROUP BY tags.id, tags.tag_name
121 ORDER BY common_Hashtag DESC
122 LIMIT 5;
123
```

common_Hashtag	tag_name	id
59	smile	21
42	beach	20
39	party	17
38	fun	13
24	concert	18

- 5) Campaign Launch: Discovered the best day for user registration, providing insights for ad scheduling.

```
.24 ##Determine the day of the week when most users register on Instagram.
.25 #Provide insights on when to schedule an ad campaign.
.26
.27 • SELECT dayname(created_at) AS day, COUNT(*) AS total FROM users
.28 GROUP BY day
.29 ORDER BY total DESC;
.30
```

day	total
Thursday	16
Sunday	16
Friday	15
Tuesday	14
Monday	14
Wednesday	13
Saturday	12

- **Investor Metrics:**

- 1) User Engagement: Calculated the average number of posts per user, offering investors a view of platform activity trends..

```

131  ##Calculate the average number of posts per user on Instagram.
132  #Also, provide the total number of photos on Instagram divided by the total number of users.
133
134  • SELECT (COUNT(photos.id) * 1.0) / COUNT(DISTINCT photos.user_id) AS avg_posts_per_user,
135         (COUNT(photos.id) * 1.0) / (SELECT COUNT(id) FROM users) AS total_photos_per_user
136         From photos;
137

```

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

avg_posts_per_user	total_photos_per_user
3.47297	2.57000

- 2) Bots & Fake Accounts: Identified potential bot accounts that have liked every post, ensuring a better understanding of fake engagement.

```

138  ##Identify users (potential bots) who have liked every single photo on the site,
139  #as this is not typically possible for a normal user.
140
141  • SELECT likes.user_id AS Potential_bot FROM likes
142         GROUP BY likes.user_id
143         HAVING COUNT(DISTINCT likes.photo_id) = (SELECT COUNT(user_id) FROM photos);
144

```

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

Potential_bot
5
14
21
24
36
41
54
57
66
71
75
76
91

Result:

This project gave us important insights that can help shape Instagram's future plans. These findings can help the marketing team create targeted ads and the product team improve user experience. Investors also get a clearer view of how genuine and active the platform is.