SUMIT SHAMLAL CHAURE

INSTAGRAM USER ANALYTICS
Trainity Project 2 – SQL Fundamentals

Figure 1 - https://metricswatch.com/instagram-analytics-report-template

Introduction

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

Description

Imagine you're a data analyst working with the product team at Instagram. Your role involves analyzing user interactions and engagement with the Instagram app to provide valuable insights that can help the business grow.
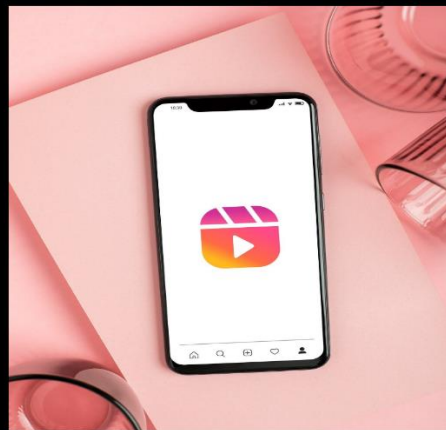
User analysis involves tracking how users engage with a digital product, such as a software application or a mobile app. The insights derived from this analysis can be used by various teams within the business. For example, the marketing team might use these insights to launch a new campaign, the product 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 analyze 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.

SS - 1) Webpick Image

Requirements -

1) Project Description :

The aim of the project is to find the user analytics of the instagram database and get insights about user activities, engagement with the platform, post frequency, traffic incoming, ad monetization planning and various management activities to grow the platform. It is our work to do a thorough analysis on the mentioned points and bring on the desired results.For handling the data we will use the given query to create a dummy database and then go through the links connection among various tables and understand the logic of basic working of the platform to query on the proper factors.

2) Approach :

Firstly I created the database and tables from the provided query and learnt about the connection among them to better query in the required output.For the visualization part I made the csv data into excel tables and generated the screenshots of both the query and table for results.

The analysis based on query has insights at the bottom of the screengrabs to let the repartee understand the aim of each analysis.

3) Tech-Stack Used :

MySQL – For the main query and results part I have relied on MySQL version 8.1.0.

Excel – The tables were generated from the csv data of the query results of sql and imported in excel.

Word – The report is written in word/docx format using MS Word and then exported to pdf.

Drive – To upload all the essential files attached in the report for reference and for pdf upload.

4) Insights :

The summary for each query is given with the screenshot but to summarize the overall thing I came to the conclusion that there are about a 100 registered user on the platform along with a good 257 post uploaded. Also there are around 7488 comments on all the posts combined with around 21 popular hashtags used across the posts and around 501 photo tags embedded. From the more broader view the insights enlightens us with the active days for helping the campaign to broadcast ads and gain more profits same with the registration days to gain more traffic and notify the teams to manage the incoming flow. The user activities can be tracked and the inactive user can be notified by the team to start posting and award the loyal and active users. This and some other insights which are really helpful from the investors and development point of view are discussed below.

5) Result :

The project has taught me about the gist of real life example of database operations from table creation, data insertion management to the complexity involved in to connect various tables to store and get relevant data. The analysis helps to know about the parameter and inputs required to track the activities to keep the site running and also on a broader level how to keep the site engaged and attract more customers to bring In the revenue for the company. The results here help us to identify suspicious as well as inactivity of users and also the points to consider during data analysis. I learnt about joins and select statements more practically with the activity.

Main Querying Part & Analysis -

SQL Task

A) Marketing Analysis:

    1. Loyal User Reward

    2. Inactive User Engagement

    3. Contest Winner Declaration

    4. Hashtag Research

    5. Ad Campaign Launch

B) Investor Metrics:

    1. User Engagement
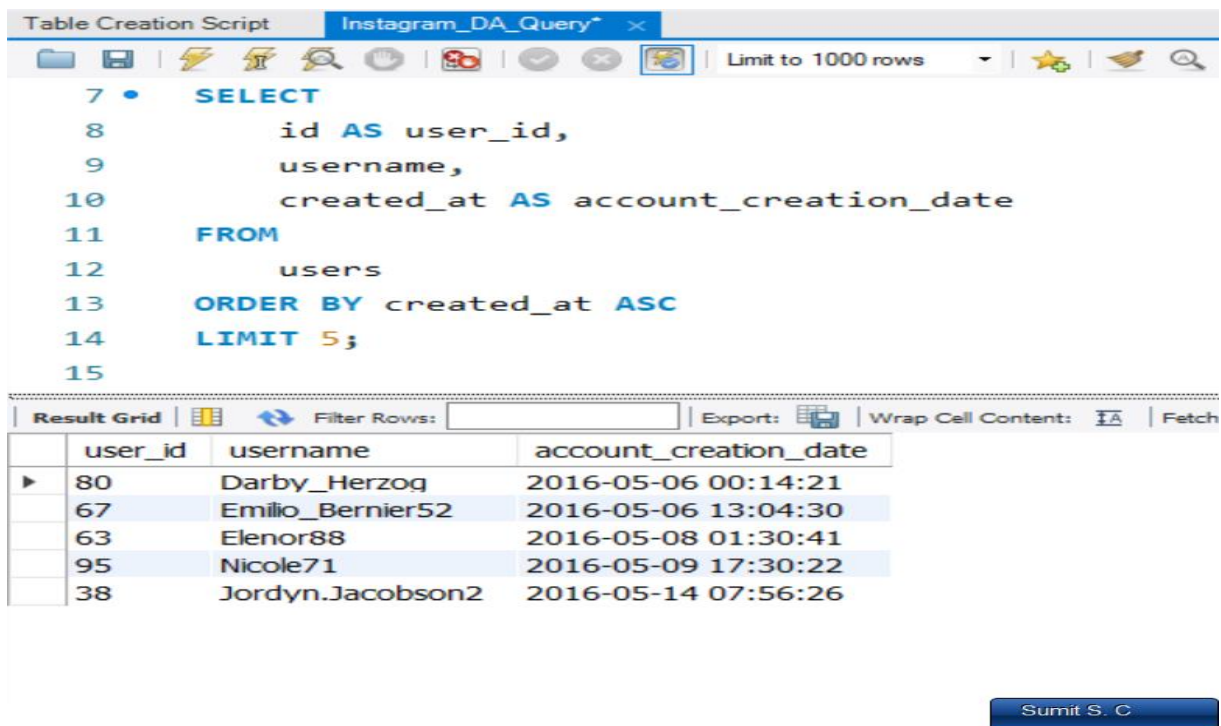
    2. Bots & Fake Accounts

A) Marketing Analysis:

Querying Related to the activities of site for user interaction & user management and to grow the traffic.

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.

Query

user_id	username	account_creation_date
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

	A	B	C	D	E
1			**Five Oldest Instagram Users**		
2		user_id	username	account_creation_date	
3		80	Darby_Herzog	06-05-2016 00:14	
4		67	Emilio_Bernier52	06-05-2016 13:04	
5		63	Elenor88	08-05-2016 01:30	
6		95	Nicole71	09-05-2016 17:30	
7		38	Jordyn.Jacobson2	14-05-2016 07:56	
8					

SS - 2) Five Oldest Users on Platform

Insights

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.

Query

```
16    -- 2)Inactive User Engagement:  Identify
17 •  SELECT
18        u.id AS user_id,
19        u.username,
20        p.id AS photo_id,
21        p.image_url AS post_url
22    FROM
23        users u
24            LEFT JOIN
25        photos p ON u.id = p.user_id
26    WHERE
27        p.user_id IS NULL;
28
```

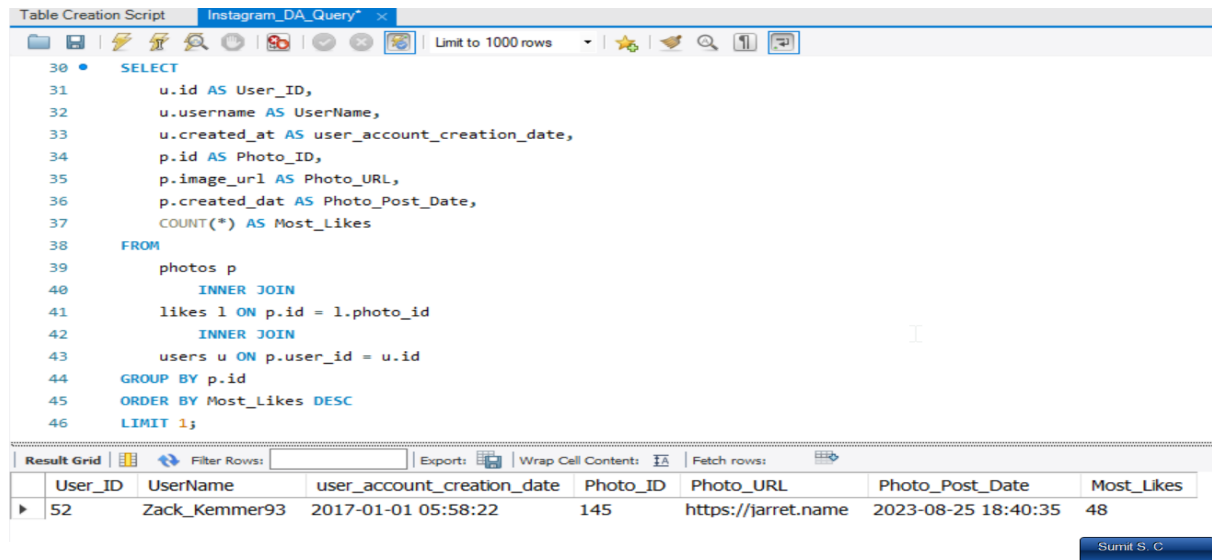| user_id	username	photo_id	post_url
Inactive User - No Single Post/Photos on Platform			
Note - Photo url & post id are blank denotes user is inactive			
5	Aniya_Hackett	NULL	NULL
7	Kasandra_Homenick	NULL	NULL
14	Jaclyn81	NULL	NULL
21	Rocio33	NULL	NULL
24	Maxwell.Halvorson	NULL	NULL
25	Tierra.Trantow	NULL	NULL
34	Pearl7	NULL	NULL
36	Ollie_Ledner37	NULL	NULL
41	Mckenna17	NULL	NULL
45	David.Osinski47	NULL	NULL
49	Morgan.Kassulke	NULL	NULL
53	Linnea59	NULL	NULL
54	Duane60	NULL	NULL
57	Julien_Schmidt	NULL	NULL
66	Mike.Auer39	NULL	NULL
68	Franco_Keebler64	NULL	NULL
71	Nia_Haag	NULL	NULL
74	Hulda.Macejkovic	NULL	NULL
75	Leslie67	NULL	NULL
76	Janelle.Nikolaus81	NULL	NULL
80	Darby_Herzog	NULL	NULL
81	Esther.Zulauf61	NULL	NULL
83	Bartholome.Bernhard	NULL	NULL
89	Jessyca_West	NULL	NULL
90	Esmeralda.Mraz57	NULL	NULL
91	Bethany20	NULL	NULL
Inactive User Count	26		

Sumit S. C

Sumit S. C

SS - 3) Inactive Users

Insights

*From The above table we can conclude that about "**26**" different individuals are currently inactive or ghost users on our platform based on the query which checked whether a user has posted at least a single post/photo from joining day.*

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.

Query

SS - 4) Most Like Post On Instagram – User Details

The insight of the query shows us that **" Zack_Kemmer93"** photo with post id – **145** posted on **25th august 2023** is the Most Liked Photo in our data.

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.

Query

```
48      -- 4)Hashtag Research: Identify and suggest the t
49 •    SELECT
50          t.tag_name,
51          COUNT(p.tag_id) AS tag_count
52      FROM
53          tags t
54      JOIN
55          photo_tags p ON t.id = p.tag_id
56      GROUP BY t.tag_name
57      ORDER BY tag_count DESC
58      LIMIT 5;
59
```

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

Top 5 Hashtags

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

SS - 5) Most Used Hashtag – Top 5

Insights

*From the analysis its evident that most users use the hashtags like – **"Smile, Beach, Party, Fun, Concert"** with their post to be interactive and stay in the feeds.*

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.

Query

```
60    -- 5)Ad Campaign Launch: Determine the day of the wee
      ad campaign.
61 •  SELECT
62        DATE_FORMAT(created_at, '%W') AS Day_Of_Week,
63        COUNT(*) AS total_registered
64    FROM
65        users
66    GROUP BY DAY_OF_week
67    ORDER BY total_registered DESC
68    LIMIT 3;
69
```

	A	B
1	**Best Day For Ad-Launch (Most Registered Day)**	
2	Day_Of_Week	total_registered
3	Thursday	16
4	Sunday	16
5	Friday	15
6		

Sumit S. C

Sumit S. C

SS - 6) Most Active day to Launch Ads

Insights

*We found from our analysis that most new users registered on "**Thursdays & Sunday**"*

*and on 2nd number is **Friday** when the Campaign teams can launch ad to grab more*

traffic and clicks.

B) Investor Metrics:

Query related to investors concerns like site performance and suspicious activities that can hamper the service or decrease the revenue.

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.

Query

```
75  •       SELECT
76  ⊖           (SELECT
77                      COUNT(*) AS total_users    -- Total users
78                  FROM
79                      users) AS Total_Users,
80  ⊖           (SELECT
81                      COUNT(*) AS total_photos    -- Total posts
82                  FROM
83                      photos) AS Total_Photos,
84  ⊖           (SELECT
85                      COUNT(*)
86                  FROM
87                      photos) / (SELECT
88                      COUNT(*)
89                  FROM
90                      users) AS avg_posts_per_user;
91
```

Result Grid | 🔢   ↔ Filter Rows: [   ] | Export: 🔣 | Wrap Cell Content:

Total_Users	Total_Photos	avg_posts_per_user
100	257	2.5700

Sumit S. C

	A	B	C
1	**User Engagement Check**		
2	Total_User ▼	Total_Photo ▼	avg_posts_per_use ▼
3	100	257	2.57
4			

Sumit S. C

SS - 7) User Engagement Check for investor Matrices
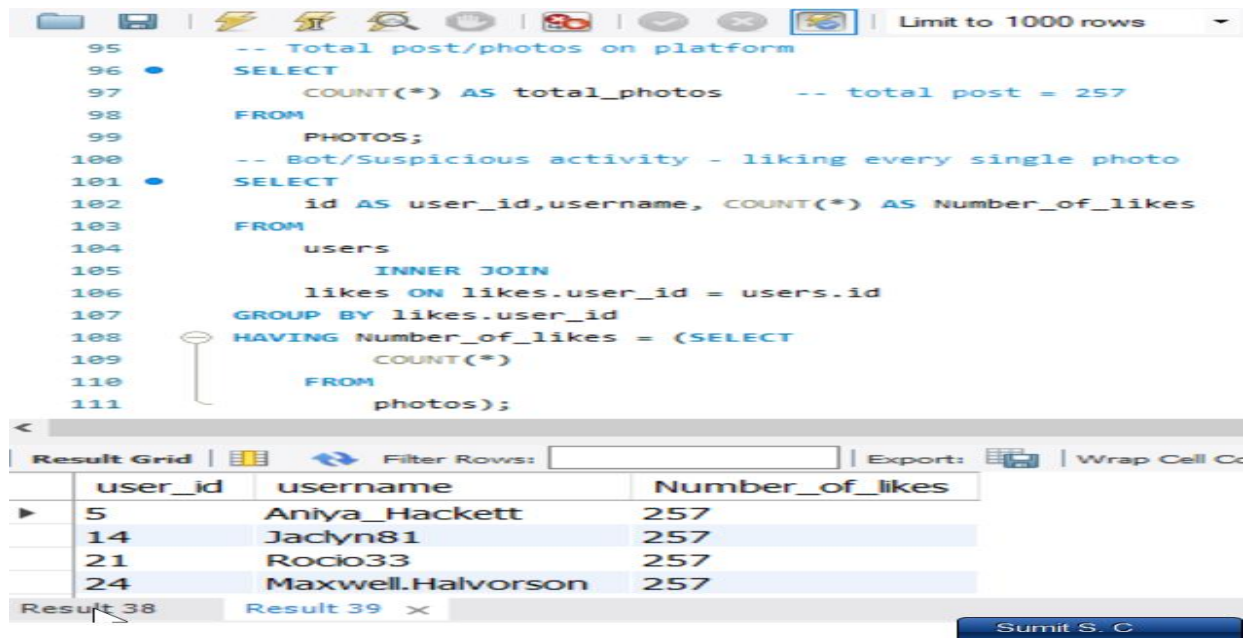
---

*The Above results show us that there are about "**100**" odd accounts on our Platform &
about "**257**" Photo posts so that averages about "**2.57 posts per user**" which shows that
there is a healthy audience on the platform for various activities like growth, ads and
revenue generation.*

---

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.

Query

	A	B	C	D
1		Bot Activity Check - Using Post Likes		
2		user_id	username	Number_of_likes
3		5	Aniya_Hackett	257
4		14	Jaclyn81	257
5		21	Rocio33	257
6		24	Maxwell.Halvorson	257
7		36	Ollie_Ledner37	257
8		41	Mckenna17	257
9		54	Duane60	257
10		57	Julien_Schmidt	257
11		66	Mike.Auer39	257
12		71	Nia_Haag	257
13		75	Leslie67	257
14		76	Janelle.Nikolaus81	257
15		91	Bethany20	257
16		Suspicious Accounts Count	13	
17		Total_Post	In Our check if the user who likes all the posts on platform can be considered as a bot	
18		257		

Sumit S. C

SS - 8) Bot Activities

Insights

*The above query has helped us identify that there is a total of "**257**" Posts/Photos on the platform and about "**100**" Users there seem to be some odd and suspicious activity by around "**13**" users as they have liked every post's which seems or can be interpreted as a bot activity.*

Note – Liking everything is not a valid parameter but in real life example it's impossible for normal user to like every photo's posted on platform so these can be taken as bots.

Drive Link – Contains files for reference like SQL query, text file , excel/csv for reference & proof.

(Download the zip file and open the files as google sheets and docs changes the formatting of tables and report files)

SS - 9) canva template closing

Thank You