Instagram User Analytics

SQL Fundamentals

Project Description : This is a project about Instagram User analysis in which I helped the attempt to derive business insights for marketing, product & development teams by providing the data to them by using SQL(structured query language) On this process am using several SQL queries to get the desired data Through the project I founded the top oldest user, whether the user posted any picture or not, most commonly used hashtag, total number of users and many several insights of the Instagram from the data provided.

Approach : I carefully understood the task and looked what actual data the team need and then first I imported the data into SQL and perform several commands to understand the data then find the insights which the team needs to perform their marketing and development.

Tech-stack Used : I used db-fiddle(MySQL v5.7) and MySQL Workbench 8 0 CE

Insights : I performed several SQL Commands to get the insights and got a knowledge of how to perform a real time SQL knowledge on this project I understood the data given of Instagram users, I derived several insights about total users of Instagram, oldest user and is there any fake account present on Instagram Understood the process of SQL.

Result : I derived the several data from the given data and the data I derived is mentioned below And this helped me to understand the real-time SQL commands.

A) Marketing :

1) Rewarding Most Loyal Users : People who have been using the platform for the longest time.

Task-1 : Find the 5 oldest users of the Instagram from the database provided?

Query : select distinct(username), created_at from ig_clone.users

order by created_at limit 5;

Output :

Darby_Herzog	2016-05-06 00:14:21
Emilio_Bernier52	2016-05-06 13:04:30
Elenor88	2016-05-08 01:30:41
Nicole71	2016-05-09 17:30:22
Jordyn.Jacobson2	2016-05-14 07:56:26

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

Task-2 : Find the users who have never posted a single photo on Instagram?

Query : SELECT u.* FROM ig_clone.users u

LEFT JOIN ig_clone.photos p ON u.id = p.user_id

WHERE p.user_id is null

order by u.username;

Output :

Query #1 Execution time: 1ms

id	username	created_at
5	Aniya_Hackett	2016-12-07 01:04:39
83	Bartholome.Bernhard	2016-11-06 02:31:23
91	Bethany20	2016-06-03 23:31:53
80	Darby_Herzog	2016-05-06 00:14:21
45	David.Osinski47	2017-02-05 21:23:37
54	Duane60	2016-12-21 04:43:38
90	Esmeralda.Mraz57	2017-03-03 11:52:27
81	Esther.Zulauf61	2017-01-14 17:02:34
68	Franco_Keebler64	2016-11-13 20:09:27

74	Hulda.Macejkovic	2017-01-25 17:17:28
14	Jaclyn81	2017-02-06 23:29:16
76	Janelle.Nikolaus81	2016-07-21 09:26:09
89	Jessyca_West	2016-09-14 23:47:05
57	Julien_Schmidt	2017-02-02 23:12:48
7	Kasandra_Homenick	2016-12-12 06:50:08
75	Leslie67	2016-09-21 05:14:01
53	Linnea59	2017-02-07 07:49:34
24	Maxwell.Halvorson	2017-04-18 02:32:44
41	Mckenna17	2016-07-17 17:25:45
66	Mike.Auer39	2016-07-01 17:36:15
49	Morgan.Kassulke	2016-10-30 12:42:31
71	Nia_Haag	2016-05-14 15:38:50
36	Ollie_Ledner37	2016-08-04 15:42:20
34	Pearl7	2016-07-08 21:42:01
21	Rocio33	2017-01-23 11:51:15
25	Tierra.Trantow	2016-10-03 12:49:21

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.

Task 3 :- Identify the winner of the contest and provide their details to the team?

Query : select ig_clone.users.username,count(*) as total_likes from ig_clone.likes

join ig_clone.photos on photos.id=ig_clone.likes.photo_id

join ig_clone.users on users.id=ig_clone.likes.photo_id

group by photos.id

order by total_likes desc

limit 1;

Output :

username	total_likes
Zack_Kemmer93	41

4) Hashtag Researching : A partner brand wants to know, which hashtags to use in the

post to reach the most people on the platform.

Task 4 :- Identify and suggest the top 5 most commonly used hashtags on the platform?

Query : select ig_clone.tags.tag_name, count(pt.photo_id) as photo_tagged from ig_clone.photo_tags pt

inner join ig_clone.tags on pt.tag_id = tags.id

group by ig_clone.tags.tag_name

order by photo_tagged desc

limit 5;

Output :

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

5) Launch AD Campaign : The team wants to know, which day would be the best day to launch ADs.

Task : What day of the week do most users register on? Provide insights on when to schedule an ad campaign.

Query : SELECT DAYNAME(created_at) AS day, COUNT(*) AS total FROM ig_clone.users

GROUP BY day

ORDER BY total DESC

limit 2;

Output：

Query #1 Execution time: 0ms

day	total
Sunday	16
Thursday	16

B) investor metrics :

1) User Engagement : Are users still as active and post on Instagram or they are making fewer posts.

Task : Provide how many times does average user posts on Instagram. Also, provide the total number of photos on Instagram/total number of users.

Query : WITH CTE AS (SELECT u.id AS userid, COUNT(p.id)

AS photoid FROM ig_clone.users u LEFT JOIN

ig_clone.photos p ON u.id = p.user_id GROUP BY u.id)

SELECT SUM(photoid) AS total_photos,

COUNT(userid) AS total_users, SUM(photoid)/COUNT(userid) AS photos_per_user FROM CTE

Output :

total_photos	total_users	photos_per_user
257	100	2.5700

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

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).

Query : WITH photo_count AS (SELECT user_id, COUNT(photo_id) AS num_like

FROM ig_clone.likes GROUP BY user_id ORDER BY num_like DESC)

SELECT * FROM photo_count WHERE num_like =

(SELECT count(*) FROM ig_clone.photos)

Output :

user_id	username	total_user_likes
5	Aniya_Hackett	257
14	Jadyn81	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