Instagram User Analytics (SQL Fundamentals)

By

Dinesh Jathar

Project Description

- This project is about finding out the various insights in Instagram User Database using numbers and values provided by raw materials in this project. We are helping our Management team and our superiors by providing insights for asked questions by them.
- For example:
- Oldest Users
- Total number of Photos
- Fake Accounts
- Top Used Hashtags
- Active Users
- Average User Posts on Instagram.

Project Approach

For proceeding of this project ,we need SQL as a tool for fetching results by framing queries on based of the asked questions. We must create the database for fetching results, once the database was created, various functions and data extracting were used to get the insights required.

Tool Used For the Project

- MY SQL Workbench was used during execution in order to query database.The ease of access and setup , troubleshooting support as well as the GUI made it a good tool for the project.
- Project Insights are as follows:

1. Marketing: The marketing team wants to launch some campaigns, and they need your help with the following,
Rewarding Most Loyal Users
People who have been using the platform for the longest time.
Task: Find the 5 oldest users of the Instagram from the database provided.

```
SELECT * FROM USERS
ORDER BY CREATED_AT
LIMIT  5;
```

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

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

```
select users.id, users.username,photos.user_id
from users
left join photos on users.id = photos.user_id
where photos.user_id is null;
```

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

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: Identify the winner of the contest and provide their details to the team

```
select photos.user_id,likes.photo_id,
count(likes.user_id) as count from ig_clone.likes
join ig_clone.photos on
ig_clone.likes.photo_id = ig_clone.photos.id
group by likes.photo_id
order by count desc
limit 1;
```

	user_id	photo_id	count
▶	52	145	48

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: Identify and suggest the top 5 most used hashtags on the platform.

```
select tags.tag_name,photo_tags.tag_id,
count(photo_tags.photo_id) as count from
ig_clone.photo_tags
join ig_clone.tags on
ig_clone.photo_tags.tag_id = ig_clone.tags.id
group by photo_tags.tag_id
order by count desc
limit 5;
```

tag_name	tag_id	count
smile	21	59
beach	20	42
party	17	39
fun	13	38
concert	18	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.

```
select dayname(created_at) as week_day_name,count(*) as total_registration
from ig_clone.users
group by week_day_name
order by total_registration desc;
```

week_day_name	total_registration
Thursday	16
Sunday	16
Friday	15
Tuesday	14
Monday	14
Wednesday	13
Saturday	12

6. 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.
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.

```
select round((select count(*) from ig_clone.photos)/(74),2) as avg_users;
```

avg_users
6.95

7. 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.
User Engagement : Are users still as active and post on Instagram or they are making fewer posts
Task: Provide the total number of photos on Instagram.

```
select count(id) as total_users
from ig_clone.users;
```

total_photos
514

8. 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.
User Engagement: Are users still as active and post on Instagram or they are making fewer posts.
Task: Provide total number of users.

```
select count(id) as total_users

from ig_clone.users;
```

total_users
100

9. 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.

```
select likes.user_id,users.username,count(likes.photo_id) as count
from ig_Clone.likes
join ig_clone.users on
ig_clone.likes.user_id = ig_clone.users.id
group by likes.user_id
order by count desc;
```

user_id	username	count
21	Rocio33	257
71	Nia_Haag	257
5	Aniya_Hackett	257
66	Mike.Auer39	257
41	Mckenna17	257
14	Jaclyn81	257
57	Julien_Schmidt	257
24	Maxwell.Halvorson	257
76	Janelle.Nikolaus81	257
75	Leslie67	257
54	Duane60	257
91	Bethany20	257
36	Ollie_Ledner37	257
16	Annalise.McKenzi...	103
96	Keenan.Schamber...	98
69	Karley_Bosco	97
65	Adelle96	96

user_id	username	count
2	Andre_Purdy85	94
26	Josianne.Friesen	94
4	Arely_Bogan63	93
13	Alexandro35	93
87	Rick29	92
35	Lennie_Hartmann40	92
93	Willie_Leuschke	91
92	Frederik_Rice	91
22	Kenneth64	91
32	Irwin.Larson	91
19	Hailee26	90
39	Kelsi26	89
11	Justina.Gaylord27	89
46	Malinda_Streich	88
62	Ressie_Stanton46	88
70	Erick5	88
31	Aiyana_Hoeger	88

user_id	username	count
10	Presley_McClure	87
42	Maya.Farrell	87
85	Milford_Gleichner42	87
20	Delpha.Kihn	87
44	Seth46	86
43	Janet.Armstrong	86
60	Sam52	86
67	Emilio_Bernier52	86
95	Nicole71	86
73	Jaylan.Lakin	86
38	Jordyn.Jacobson2	85
9	Gus93	85
52	Zack_Kemmer93	85
72	Kathryn80	85
40	Rafael.Hickle2	85
15	Billy52	84
37	Yazmin_Mills95	84

user_id	username	count
82	Aracely.Johnston98	84
94	Damon35	84
61	Jayson65	83
63	Elenor88	83
78	Colten.Harris76	83
6	Travon.Waters	82
100	Javonte83	82
18	Odessa2	82
30	Kaley9	81
50	Gerard79	81
56	Peter.Stehr0	81
8	Tabitha_Schambe...	79
3	Harley_Lind18	79
27	Darwin29	79
17	Norbert_Carroll35	78
55	Meggie_Doyle	78
33	Yvette.Gottlieb91	77

12	Dereck65	77
28	Dario77	77
47	Harrison.Beatty50	76
48	Granville_Kutch	75
79	Katarina.Dibbert	75
84	Alysa22	75
98	Imani_Nicolas17	74
99	Alek_Watsica	74
97	Tomas.Beatty93	69

Conclusion

Through working on this project, I have gained lot of insights like oldest users, Total number of photos, Average user posts, Commonly used hashtags, Winner of the contest, Fake users or many more. My SQL Workbench and PowerPoint are the tools used in this project; this project improves my presentation skills as well as enhance my SQL knowledge.

THANK YOU