Instagram User Analytics

Bandi Chakradhar

Aims

This project is aimed at finding certain marketing metrics and investor metrics from Instagram database.

Approach

Firstly, we created an SQL database with provided data. We closely studied the requirements of the projects and metrics to be delivered for various teams.

Used MySQL workbench to query database and produce valuable insights into Instagram user base.

Tech Stack

We used MySQL Workbench 8.0 CE to create and extract useful insights from Instagram database.

Insights and Results

We present the each metrics with MySQL code as Follows:

A) Marketing Metrics

1.Rewarding the most loyal users: People who have been using the platform for the longest time.(Top 5 oldest Instagram users)
Here, We will use the data from the users table by selecting the username and created_at columns.
Query:

```
select username, created_at
from users
order by created_at ASC
limit 5;
```

Output :
Oldest 5 users with their created_at date and time are as follows

username	created_at
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: Remind Inactive users to Start Posting(Users who never posted a single photo on Instagram)

Here, we will collect the id and username of user from users table for which the id is not present in the list of userids from photos table.

Query:

```
select id,username
from users );
where  id NOT IN (select user_id from photos);
```

Output;

There are total 26 users who never posted a photo on the platform

id	username
5	Aniya_Hackett
7	Kasandra_Homenick
14	Jaclyn81
21	Rocio33
24	Maxwell.Halvorson
25	Tierra.Trantow
34	Pearl7
36	Ollie_Ledner37
41	Mckenna17
45	David.Osinski47
49	Morgan.Kassulke
53	Linnea59
54	Duane60
57	Julien_Schmidt
66	Mike.Auer39
68	Franco_Keebler64
71	Nia_Haag
74	Hulda.Macejkovic

75	Leslie67
76	Janelle.Nikolaus81
80	Darby_Herzog
81	Esther.Zulauf61
83	Bartholome.Bernhard
id	**username**
89	Jessyca_West
90	Esmeralda.Mraz57
91	Bethany20

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 and declare as the winner.

Here, first we will select the userid, username, photosid and count of likes , and then we will use innerjoin with three tables photos, likes, and users and using groupby on basis of photos.id.

Query:

```
select users.id as user_id, users.username, photos.id as photo_id, count(*) as nlikes
from photos
inner join likes
on likes.photo_id = photos.id
inner join users
on photos.user_id = users.id
group by photos.id
order by total DESC
limit 1;
```

Output:

User_id	username	Photo _id	nlikes
52	Zack_Kemmer93	145	48

So, the user named Zack_Kemmer93 with user_id 52 is the winner of the contest cause his photo with photo_id 145 has the highest number of likes i.e. 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.(Top 5 commonly used #Hashtags on Instagram)

Here, we need **t**o join tags table and photo_tags table, **on tags.id = photo_tags.tag_id** and then by using group by function we need to group the desired output on the basis of **tags.tag_name**

Query:

```
select tags.tag_id, tags.tag_name, count(*) as ntimes_of_tag
from tags join photo_tags
on tags.id = photo_tags.tag_id
group by tags.tag_name
order by total_number_of_times_tag_used_individually DESC
limit 7;
```

Output:

tag_id	tag_name	ntimes_of_tag
21	smile	59
20	beach	42
17	party	39
13	fun	38
5	food	24
11	lol	24
18	concert	24

5.Launch AD Campaign : The team wants to know, which day would be the best day to launch ADs. (What day of the week do most users register on)

Query:

```
select dayname(created_at) as day_week, count(*) as total_number_of_users_registered
from users
group by day_week
order by total_number_of_users_registered DESC;
```

Output:

day_week	Total_number_of-users_registered
Thursday	16
Sunday	16
Friday	15
Tuesday	14
Monday	14
Wednesday	13
Saturday	12

Thursdays and Sundays are two days with most user registration on Instagram.

B) Investor Metrics;

1.User Engagement : How many times does average user posts on Instagram? Also, provide the total number of photos on Instagram/total number of users.

Total number of photos / Total number of users

Query:

```
select (select count(*) from photos)/(select count(*) from users) as
total_avg_photos;
```

On an average a user posted 2.57 photos on Instagram.

```
select user_id,count(*) as user_post_count
from photos
group by user_id
order by user_id;
```

Output;

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

Here, we use inner join to users and likes table on the basis of users.id and likes.user_id, using the on function/clause and then by using the group by function we group the desired output table on the basis of likes.user_id .

Query:

```
select user_id, username, count(*) as total_likes
from users
```

inner join likes
on users.id = likes.user_id
group by likes.user_id having total_likes = (select count(*) from photos);

Output:

user_id	username	total_likes
5	Aniya_Hackett	257
14	Jaclyn81	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