Instagram User Analytics

Nikhil Panda

Project Overview

This project is focused on analyzing the user data in the Instagram-like database.
User analysis is the process by which we track how users engage and interact with our digital product (software or mobile application) in an attempt to derive business insights for marketing, product & development teams.

Approach & Tech stack used

An Instagram identical database was created with likes, comments, photos, tags and other required data respective to the features of our software application

MySQL Workbench CE was used for completing this project.

Marketing Insights

The marketing team wants to launch some campaigns.
These Insights can be used by the marketing team to launch new Campaigns, decide the winners of contests, rewards etc.. also used to decide on features to build for an app. track user engagement and improve the experience altogether while helping the business grow.

Rewarding Most Loyal Users: 5 People who have been using the platform for the longest time.

Most loyal users are the people who have been using Instagram for the longest time. We can find this by analyzing the dates the profiles were created on and setting the limit to 5, which returns only 5 entries.

```
-- A.1) rewarding most loyal users

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

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

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

Inactive Users are users who have never posted a single post on Instagram. This can be done by checking the user entries with no photos. There are 26 users with no posts.

```sql
-- A.2) remind inactive users to start posting
select username, photos.id as posts
from users
left join photos on users.id = photos.user_id
where photos.id is null;
```

username	posts
Aniya_Hackett	NULL
Bartholome.Bernhard	NULL
Bethany20	NULL
Darby_Herzog	NULL
David.Osinski47	NULL
Duane60	NULL
Esmeralda.Mraz57	NULL
Esther.Zulauf61	NULL
Franco_Keebler64	NULL
Hulda.Macejkovic	NULL
Jadyn81	NULL

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.

Likes per photo can be found using the users. photos and likes table to find the photo that belongs to a single user with most likes.
Joining the three tables to finding and declaring the contest winner.

```
-- A.3) declaring contest winner
select username, photos.id as ID,count(*) as Total_Likes
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_Likes desc
limit 1;
```

	Contest Winner	ID	Total_Likes
▶	Zack_Kemmer93	145	48

Hashtag Researching: A partner brand wants to know, which hashtags to use in the post to reach the most people on the platform.

To find the 5 most used hashtags on Instagram. This can be done by number of times the tags has been used from the photo tags table and grouping by tags while joining the tags and photo_tags tables and setting the limit to 5.

```sql
-- A.4) hashtag researching
select tags.tag_name, count(*) as Total_tags
from photo_tags
join tags on tags.id = photo_tags.tag_id
group by tags.id
order by Total_tags desc
limit 5;
```

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

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

Using the created_at data from the users table, converting the date to day using the DAYNAME() aggregation function that returns what day of the week it was on the specified date. Then, grouping the data based on the days of the week in descending order to find out the most active day of the week.

```sql
-- A.5) Launch AD Campaign
select dayname(created_at) as day, count(*) as Total_reg
from users
group by day
order by Total_reg desc
limit 1;
```

	day	Total_reg
▶	Thursday	16

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.

Provide how many times does average user posts on Instagram. The total number of photos on Instagram/total number of users is the Average posts per person on Instagram.

```
-- B.1) User Engagement
select round(
(select count(*) from photos) /
(select count(*) from users))
as `Average Posts per User`;
```

Average Posts per User
3

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

Required to 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). Selecting the usernames with number of likes equal to the number of photos.

```sql
-- B.2) Bots & Fake accounts
select id, username, count(*) as total_likes
from users
join likes on likes.user_id = users.id
group by users.id
having total_likes = (select count(*) from photos);
```

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

THANK YOU!