INSTAGRAM USER ANALYTICS

Description:

User Analytics helps the organization to learn about user activities related to the application/ product and helps in getting insights from it to improve the user experience.
In the project we have been given Instagram user data which we need to analyze to find solutions to the problems of the Instagram management team regarding marketing and investing.

Approach:

Firstly, we needed to understand the data and the features that were available to us. We also needed to understand the questions exactly to write appropriate SQL queries in order to get desired results.

Tech Stack Used:

MySQL server version 8.0.32 was used to execute the project. It provides faster results just with the use of simple queries.

Project Insights:

A) MARKETING:

1. REWARDING MOST LOYAL USERS

QUERY: Select * from users order by created_at limit 5;

```
mysql> 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 |
+----+-------------------+---------------------+
5 rows in set (0.00 sec)
```

INSIGHT: The users with user id (80,67,63,95,38) are the 5 oldest users of the app.

2. REMINDING INACTIVE USERS TO START POSTING

QUERY: Select id,username from users where id not in (Select user_id from photos);

```
mysql> select id,username from users where id not in (select user_id from photos);
+----+--------------------+
| 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 |
| 89 | Jessyca_West       |
| 90 | Esmeralda.Mraz57   |
| 91 | Bethany20          |
+----+--------------------+
26 rows in set (0.00 sec)
```

INSIGHT: The users with user id (5,7,14,21,24,25,34,36,41,45,49,53,54,57,66,68,71,74,75,76,80,81,83,89,90,91) haven't posted anything. So, giving them reminders can help. Emailing them regarding new features and friends might help here.

3. DECLARING CONTEST WINNER (having most likes on single pic)

QUERY: select username, photos.id, photos.image_url, count(*) as total 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;

```
mysql> select username, photos.id, photos.image_url, count(*) as total from photos inner join likes on likes.photo_id=ph
otos.id inner join users on photos.user_id=users.id group by photos.id order by total desc limit 1;
+--------------+-----+--------------------+-------+
| username     | id  | image_url          | total |
+--------------+-----+--------------------+-------+
| Zack_Kemmer93 | 145 | https://jarret.name |    48 |
+--------------+-----+--------------------+-------+
```

INSIGHT: The user with user id (52) has got the most likes on photo with id (145).

4. HASHTAG RESEARCHING (5 most used tag)

QUERY: Select tag_id , count(tag_id) from photo_tags group by tag_id order by count(tag_id) desc limit 5;

```
mysql> select tag_id, count(tag_id) from photo_tags group by tag_id order by count(ta
g_id) desc limit 5;
+--------+---------------+
| tag_id | count(tag_id) |
+--------+---------------+
|     21 |            59 |
|     20 |            42 |
|     17 |            39 |
|     13 |            38 |
|     18 |            24 |
+--------+---------------+
5 rows in set (0.00 sec)
```

QUERY: Select id, tag_name from tags where id in (21,20,18,17,13);

```
mysql> select id, tag_name from tags where id in (21,20,18,17,13);
+----+----------+
| id | tag_name |
+----+----------+
| 13 | fun      |
| 17 | party    |
| 18 | concert  |
| 20 | beach    |
| 21 | smile    |
+----+----------+
5 rows in set (0.00 sec)

mysql>
```

INSIGHT: #fun , #party , #concert ,#beach , #smile are the most used tags. Using these tags will bring more engagement and if any business or ads will use these hashtags they will get more audience.

5. **LAUNCH AD CAMPAIGN (best day for launching campaign)**

QUERY:select dayname(created_at),count(dayname(created_at)) from users group by dayname(created_at) order by count(created_at) desc ;

```
mysql> select dayname(created_at),count(dayname(created_at)) from users group by dayname(created_at);
+--------------------+----------------------------+
| dayname(created_at) | count(dayname(created_at)) |
+--------------------+----------------------------+
| Thursday           |                         16 |
| Sunday             |                         16 |
| Tuesday            |                         14 |
| Saturday           |                         12 |
| Wednesday          |                         13 |
| Monday             |                         14 |
| Friday             |                         15 |
+--------------------+----------------------------+
7 rows in set (0.00 sec)
```

INSIGHT: Thursdays and Sundays are the best days for campaigns as most new users are added on these days. So, launching an ad campaign on those days will bring more attention to it even from new users.

B) INVESTOR METRICS

1. **USER ENGAGEMENT (average number of posts per user)**

QUERY: select round((Select count(*) from photos)/select count(*) from users),2);

```
mysql> SELECT ROUND((SELECT COUNT(*)FROM photos)/(SELECT COUNT(*) FROM users),2);
+-----------------------------------------------------------------+
| ROUND((SELECT COUNT(*)FROM photos)/(SELECT COUNT(*) FROM users),2) |
+-----------------------------------------------------------------+
|                                                            2.57 |
+-----------------------------------------------------------------+
1 row in set (0.01 sec)
```

INSIGHT: The average user engagement is 2.57. Any new competition might help in increasing this, such as posting a picture with a particular theme.

2. **BOTS AND FAKE ACCOUNTS (who have likes all posts)**

QUERY: select users.id, username,count(users.id) as total_likes from users join likes on users.id=likes.user_id group by users.id having total_likes =(select count(*) from photos);

```
mysql> select users.id, username,count(users.id) as total_likes from users join likes on users.id=likes.user_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 |
+----+--------------------+-------------+
13 rows in set (0.01 sec)
```

INSIGHT: The users with id(5,14,21,24,3,41,54,57,66,71,75,76,91) are bots.

Result:

Through this project we found that there are several fake accounts present on the application and the user engagement is not as high as thought. Even the maximum number of likes is not that high too. So, it has helped us to understand that steps need to be taken to increase user addition and engagement. Advertisements and introducing new features can really help such as polls, competitions which include some sort of activity.

This project has helped me to understand how different data points can be of huge relevance for a business and has helped me to work on my SQL skills.