Instagram User Analytics

Project Description: In this project, we have performed Instagram User Analytics. User analytics is a process for evaluating data that depicts how users adopt, throw themselves into, and experience products and services. This data can also be merged with user-level attributes and feedback for further decomposition. Because retaining an existing user is less expensive than having to find and convert a new one, user analytics can be used to find meaning in all of the data that your users are generating within your platform and highlight the key drivers behind what leads to attrition versus the features that keep set your product apart from the competition. The project is all about tracking user activity like how a user interacts, manages their account, and interconnects with the posts updated by other users. Then this gathered data is used by the product team and product manager across the business to decide on several features for an app, track some success factors by estimating the data extracted by performing user analytics using SQL (Structured Query Language) and also improve the user experience for better user feedback which leads in the growth of a business.

All the data extracted has been used by Marketing Teams and Investor Metrics to magnify the business process and to provide users with a mark experience through the Instagram app. We have found out the oldest users of Instagram, users who have never posted a single photo, winners of the contest i.e. the one who gets the most likes on a single photo, the top 5 popular (most commonly used) hashtags, on which day do most users register on the platform. The data pulled out from the above activity is used by the marketing teams for launching some campaigns on the platform. We have also uprooted how many times does average user posts on Instagram and provided the total number of photos on Instagram/total number of users along with data on users (bots) who have liked every single photo on the site.

Approach: The report is all about user interaction with the Instagram app, the extracted data of the users sure going to help them as well as a management team to improve and add features to the app. By using a simple approach, with the help of SQL queries the data needed for the management team is smoothly drawn out and the output is shown in the result section below.

Tech-Stack Used: For providing the information, operations are performed by installing MySQL 8.0.28 Server and MySQL Workbench 8.0.31. The provided dataset is loaded in MySQL Workbench 8.0.31 which helped me to execute queries more efficiently, also MySQL Workbench 8.0.31 provides a superb view of databases. After connecting the CMD (Command Prompt) to the MySQL 8.0.28 Server, SQL queries are performed in Command Prompt platform.

Insights: Based on the SQL operations performed on the dataset, we get to know the user's involvement with the Instagram functions and features, the drawn out data describes the need for improvement in the app and helps the management team to enhance users' experience through the app and get to know how to catch the eye of users constantly towards the app.

Initially, we rewarded the top 5 oldest users (**Darby_Herzog, Emelio_Berneir52, Elenor88, Nicole71, Jordyn.Jacobson2**) who were using Instagram for the longest time as the most loyal users. These users were constantly in touch with the app since over-long time. If a user had registered on Instagram handle for a long time but never posted a single photo on the Instagram handle, I have enlisted some users from the dataset who have never posted a single photo on Instagram for sending them an email to remind them to post their 1st photo on the platform. This can help the marketing team to catch up the attention of the users towards the app after receiving mail from the Instagram team, he/she gets ready to post a photo on the Instagram platform. Then I have taken out the data of the user who gets the highest number of likes on a single photo and declared him/her as a winner of the contest. **David.Osinski47** with the most likes (48) on a single photo is the winner of the contest. This narrates that the user (contest winner) has followers in excellent numbers and his/her interaction with the Instagram handle is often good. After recognizing the top 5 most commonly used hashtags (#smile, #beach, #party, #fun, #concert) by users, I get to know users are using these tags in large numbers to get likes on their posts in huge numbers as well to increase their follower too. Identified most user registers on **Thursday** and **Sunday** as compared to other days of the week these measures helped the marketing team to promote advertisements of brands or to schedule any campaign on these days as most of the user remains active on these two days of the week. The as large number of users were able to discern the promoted ads on these days and the marketing team will focus more on these couple of days for scheduling the campaign and promoting advertisements to gain users' attraction in massive numbers. The above data is more useful to the marketing team to track how users engage and interact with their software, product, or application in an attempt to improve their product, bring more users in, improve user engagement with their product, and the general success of their application.

Also provided data to Investors concerning the total number of users registered and how many times does average user posts on Instagram **(2.57)**. In addition, the report also contains data about fake and dummy accounts whether an account is handled by a real person or a bot. These data might help the investors to judge whether their financing in Instagram is profitable or not and gave them a glimpse of users' collaboration on the platform

Result: The result section of the report shows the answers to the questions asked by the management team.

A) Marketing: The marketing team wants to launch some campaigns, and they need your help with the following

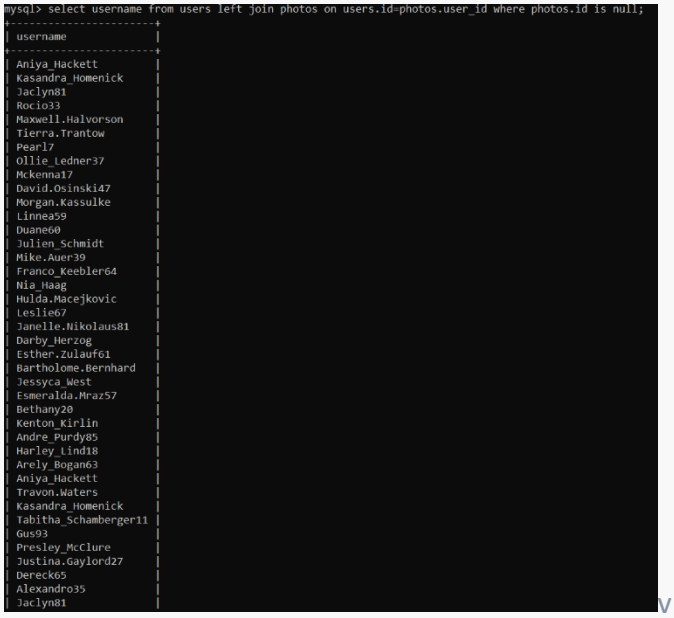
1. Rewarding Most Loyal Users: People who have been using the platform for the longest time.
Your Task: Find the 5 oldest users of Instagram from the database provided.

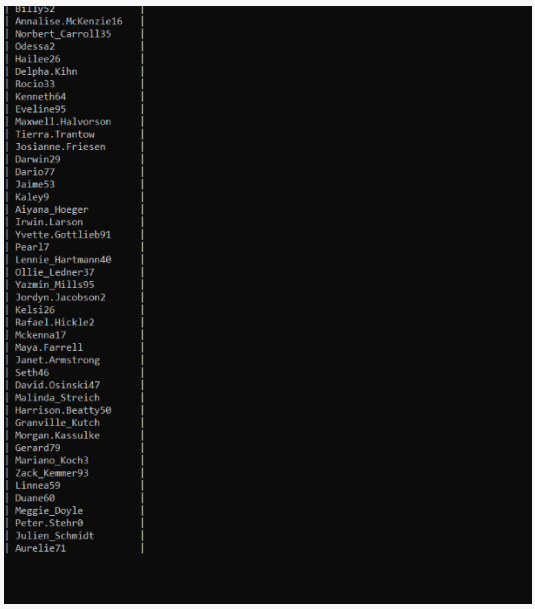
The solution for the above problem is:

```
mysql> SELECT DISTINCT(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 |
+------------------+---------------------+
5 rows in set (0.01 sec)
```

2. Remind Inactive Users to Start Posting: By sending them promotional emails to post their 1st photo.
Your Task: Find the users who have never posted a single photo on Instagram.

The solution for the above problem is:





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

The solution for the above problem is:

```
mysql> select users.username, photos.id,photos.image_url,count(*) as total_likes from ig_clone.likes join photos on photos.id=likes.photo_id join users on users.id=likes.photo_id group by p
hotos.id order by total_likes desc limit 10;
+--------------------+-----+----------------------+-------------+
| username           | id  | image_url            | total_likes |
+--------------------+-----+----------------------+-------------+
| David.Osinski47    | 145 | https://jarret.name  |          48 |
| Aracely.Johnston98 | 182 | https://dorcas.biz   |          43 |
| Darwin29           | 127 | https://celestine.name |          43 |
| Eveline95          | 123 | http://shannon.org   |          42 |
| Kaley9             |  30 | http://kenny.com     |          41 |
| Hulda.Macejkovic   | 174 | https://delbert.net  |          41 |
| Frederik_Rice      | 192 | https://anahi.info   |          41 |
| Harrison.Beatty50  | 147 | https://adela.com    |          41 |
| Jayson65           |  61 | https://dejon.name   |          41 |
| Zack_Kemmer93      |  52 | https://hershel.com  |          41 |
+--------------------+-----+----------------------+-------------+
10 rows in set (0.03 sec)
```

4. Hashtag Researching: A partner brand wants to know, which hashtags to use in the post to reach the most people on the platform.
Your Task: Identify and suggest the top 5 most commonly used hashtags on the platform.

The solution for the above problem is:

```
mysql> SELECT * FROM ig_clone.tags right join (SELECT tag_id,count(*) as number_of_times_used FROM ig_clone.photo_tags group by tag_id)t on tags.id=t.tag_id order by number
_of_times_used DESC limit 5;
+------+----------+---------------------+--------+----------------------+
| id   | tag_name | created_at          | tag_id | number_of_times_used |
+------+----------+---------------------+--------+----------------------+
|   21 | smile    | 2022-12-07 10:33:12 |     21 |                   59 |
|   20 | beach    | 2022-12-07 10:33:12 |     20 |                   42 |
|   17 | party    | 2022-12-07 10:33:12 |     17 |                   39 |
|   13 | fun      | 2022-12-07 10:33:12 |     13 |                   38 |
|   18 | concert  | 2022-12-07 10:33:12 |     18 |                   24 |
+------+----------+---------------------+--------+----------------------+
5 rows in set (0.00 sec)
```

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

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

The solution for the above problem is:

```
mysql> select dayname(created_at) as registered_day,count(*) as no_of_users from ig_clone.users group by dayname(created_at);
+----------------+-------------+
| registered_day | no_of_users |
+----------------+-------------+
| Thursday       |          48 |
| Sunday         |          48 |
| Tuesday        |          42 |
| Saturday       |          36 |
| Wednesday      |          39 |
| Monday         |          42 |
| Friday         |          45 |
+----------------+-------------+
7 rows in set (0.01 sec)
```

B) 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

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

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

The solution for the above problem is:

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

Above is the average posts by users viz. 2.57.

Below result shows number of posts by a single user:-

```
mysql> SELECT users.username,COUNT(photos.image_url) as no_of_posts FROM ig_clone.users JOIN photos ON users.id = photos.user_id GROUP BY users.id ORDER BY no_of_posts DESC;
+----------------------+-------------+
| username             | no_of_posts |
+----------------------+-------------+
| Eveline95            |          36 |
| Clint27              |          33 |
| Cesar93              |          30 |
| Delfina_VonRueden68  |          27 |
| Aurelie71            |          24 |
| Jaime53              |          24 |
| Donald.Fritsch       |          18 |
| Yvette.Gottlieb91    |          15 |
| Zack_Kemmer93        |          15 |
| Harrison.Beatty50    |          15 |
| Travon.Waters        |          15 |
| Alexandro35          |          15 |
| Mariano_Koch3        |          15 |
| Colten.Harris76      |          15 |
| Justina.Gaylord27    |          15 |
| Kenton_Kirlin        |          15 |
| Kathryn80            |          15 |
| Adelle96             |          15 |
| Janet.Armstrong      |          15 |
| Florence99           |          15 |
| Josianne.Friesen     |          15 |
| Andre_Purdy85        |          12 |
| Harley_Lind18        |          12 |
| Gus93                |          12 |
| Tabitha_Schamberger11|          12 |
| Malinda_Streich      |          12 |
| Dereck65             |          12 |
| Seth46               |          12 |
| Elenor88             |          12 |
| Irwin.Larson         |          12 |
| Dario77              |          12 |
| Annalise.McKenzie16  |          12 |
| Billy52              |          12 |
| Rick29               |          12 |
| Arely_Bogan63        |           9 |
| Presley_McClure      |           9 |
| Gerard79             |           9 |
| Emilio_Bernier52     |           9 |
| Norbert_Carroll35    |           9 |
| Norbert_Carroll35    |           9 |
| Maya.Farrell         |           9 |
| Frederik_Rice        |           9 |
| Keenan.Schamberger60 |           9 |
| Alek_Watsica         |           9 |
| Jordyn.Jacobson2     |           6 |
| Javonte83            |           6 |
| Aracely.Johnston98   |           6 |
| Alysa22              |           6 |
| Milford_Gleichner42  |           6 |
| Sam52                |           6 |
| Ressie_Stanton46     |           6 |
| Hailee26             |           6 |
| Willie_Leuschke      |           6 |
| Nicole71             |           6 |
| Kaley9               |           6 |
| Tomas.Beatty93       |           6 |
| Lennie_Hartmann40    |           6 |
| Aiyana_Hoeger        |           3 |
| Darwin29             |           3 |
| Katarina.Dibbert     |           3 |
| Rafael.Hickle2       |           3 |
| Jayson65             |           3 |
| Kelsi26              |           3 |
| Kenneth64            |           3 |
| Delpha.Kihn          |           3 |
| Granville_Kutch      |           3 |
| Odessa2              |           3 |
| Yazmin_Mills95       |           3 |
| Damon35              |           3 |
| Karley_Bosco         |           3 |
| Erick5               |           3 |
| Meggie_Doyle         |           3 |
| Imani_Nicolas17      |           3 |
| Peter.Stehr0         |           3 |
| Jaylan.Lakin         |           3 |
+----------------------+-------------+
74 rows in set (0.00 sec)
```

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

The solution for the above problem is:

```
mysql> SELECT count(distinct(photo_id)) as total_images_uploaded FROM ig_clone.likes;
+-----------------------+
| total_images_uploaded |
+-----------------------+
|                   257 |
+-----------------------+
1 row in set (0.00 sec)
```

As per the above query, the total number of photos uploaded on Instagram is **257.**
So to find out the bots mean those users who have liked every single uploaded photo which is a difficult task for a legit user, I have executed a query that shows details of the users(bots) who have liked **257(all)** uploaded photos. Details are mentioned below:-

```
mysql> select users.id,users.username,users.created_at,count(likes.photo_id) as no_of_liked_photos from users join likes on users.id=likes.user_id group by users.id,users.u
sername,users.created_at having count(likes.photo_id)=257;
+----+-------------------+---------------------+--------------------+
| id | username          | created_at          | no_of_liked_photos |
+----+-------------------+---------------------+--------------------+
|  5 | Aniya_Hackett     | 2016-12-07 01:04:39 |                257 |
| 14 | Jaclyn81          | 2017-02-06 23:29:16 |                257 |
| 21 | Rocio33           | 2017-01-23 11:51:15 |                257 |
| 24 | Maxwell.Halvorson | 2017-04-18 02:32:44 |                257 |
| 36 | Ollie_Ledner37    | 2016-08-04 15:42:20 |                257 |
| 41 | Mckenna17         | 2016-07-17 17:25:45 |                257 |
| 54 | Duane60           | 2016-12-21 04:43:38 |                257 |
| 57 | Julien_Schmidt    | 2017-02-02 23:12:48 |                257 |
| 66 | Mike.Auer39       | 2016-07-01 17:36:15 |                257 |
| 71 | Nia_Haag          | 2016-05-14 15:38:50 |                257 |
| 75 | Leslie67          | 2016-09-21 05:14:01 |                257 |
| 76 | Janelle.Nikolaus81 | 2016-07-21 09:26:09 |                257 |
| 91 | Bethany20         | 2016-06-03 23:31:53 |                257 |
+----+-------------------+---------------------+--------------------+
13 rows in set (0.04 sec)
```