Instagram User Analytics

Project Description:

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.

These insights are then used by teams across the business to launch a new marketing campaign, decide on features to build for an app, track the success of the app by measuring user engagement and improve the experience altogether while helping the business grow.

We are working with the product team of Instagram and the product manager has asked us to provide insights on the questions asked by the management team.

APPROACH:

First, identify all the necessary tables for the query. Then, join the tables that contain the data required for display or for the conditions specified in the WHERE clause. After joining the tables, display all relevant data to ensure the joins are correct and to observe the outcome of the query. This iterative process allows for optimization and validation of the query structure.

TECH STACK USED :

only used the MYSQL Workbench to clean the data, there were no duplicates and other unknown values. I also created a relational schema and executed some SQL queries on the software to find insights.

INSIGHTS: You are required to provide a detailed report answering the questions below :

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.

Task: Find the 5 oldest users of Instagram from the database provided

Q1.

```
81
82    /*Q1.) Identify the five oldest users on Instagram from the provided database.*/
83    select * from users;
84    select * from users order by created_at asc limit 5;
85
```

Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content:

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
NULL	NULL	NULL

Import records from

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

Out of the 100 users, 74 actively post pictures on Instagram, while 26 have never shared a single photo on the platform. We aim to encourage these 26 users to post their first photo by sending them promotional emails.

Q2.

```
85
86      /*Q2.) Inactive User Engagement: user who have not shown no activity in instagram.*/
87 •    select username from users left join photos on users.id = photos.user_id
88      where photos.id is null;
89
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

username
Aniya_Hackett
Kasandra_Homenick
Jadyn81
Rocio33
Maxwell.Halvorson
Tierra.Trantow
Pearl7

Result 3

3. Declaring Contest Winner:

The team initiated a contest where the user with the highest number of likes on a single photo would emerge as the winner. Now, the task is to determine the contest winner and furnish their details to the team.

Q3.

```
91
92 •    SELECT u.id, u.username, l.photo_id, COUNT(l.user_id) AS max_likes
93      FROM likes l
94      JOIN photos p ON l.photo_id = p.id
95      JOIN users u ON p.user_id = u.id
96      GROUP BY l.photo_id
97      ORDER BY max_likes DESC
98      LIMIT 1;
99
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

id	username	photo_id	max_likes
52	Zack_Kemmer93	145	48

4. Hashtag Researching:

A partner brand seeks guidance on which hashtags to use in their post to maximize reach on the platform.

Task is to identify and recommend the top five most commonly used hashtags.

Q4.

```
100  /* Q4.) Hashtag research: Identify and suggest the top five most commonly used hashtags on the platform.*/
101
102  select tags.id, tags.tag_name, count(tag_id) from tags
103  join photo_tags on tags.id = photo_tags.tag_id
104  group by tag_id
105  order by count(tag_id) desc limit 5;
106
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

id	tag_name	count(tag_id)
21	smile	59
20	beach	42
17	party	39
13	fun	38
18	concert	24

5. Launch AD Campaign:

The team seeks guidance on the optimal day to launch ads. Your task is to analyze the data to identify the day of the week with the highest number of user registrations, providing insights for scheduling the ad campaign.

```
07  /* Q5.) Ad Campaign Launch:  Determine the day of the week when most users register on Instagram.
08   Provide insights on when to schedule an ad campaign. */
09
10  select dayname(created_at),count(*) from users
11  group by dayname(created_at)
12  order by count(*) desc;
13
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

dayname(created_at)	count(*)
Thursday	16
Sunday	16
Friday	15
Tuesday	14
Monday	14
Wednesday	13
Saturday	12

B) Investor Metrics:

Our investors seek an assessment of Instagram's performance, particularly in comparison to platforms like Facebook. They wish to evaluate user engagement, specifically whether users remain active and continue to post on Instagram or if there's a decline in posting frequency. Your task is to provide the average number of posts per user on Instagram, as well as the ratio of total photos on Instagram to the total number of users.

```
14      -- ---------------------------------------------------------------------
15      -- B) Investor Metrics:
16
17      /*- User Engagement:  Investors want to know if users are still active and posting on Instagram or if they are making fewer posts.
18      Task: Calculate the average number of posts per user on Instagram. Also, provide the total
19      number of photos on Instagram divided by the total number of users. */
20
21      -- 1. Total number of posts per user
22      select count(id) from photos;
23
```

esult Grid | Filter Rows: | Export: | Wrap Cell Content:

count(id)
257

```
117     /*- User Engagement:  Investors want to know if users are still active and posting on Instagram or if they are making fewer posts.
118     Task: Calculate the average number of posts per user on Instagram. Also, provide the total
119     number of photos on Instagram divided by the total number of users. */
120
121     -- 1. Total number of posts per user
122     select count(id) from photos;
123
124     -- 2. total number of users
125     select count(id) from users;
126
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

count(id)
100

```
126
127     -- 3.  Averege user post on instagram
128
129     select avg(cnt) as avereage
130     from (select user_id, count(*) as cnt from photos
131     group by user_id
132     order by cnt desc) sub;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

avereage
3.4730

2. Bots & Fake Accounts:

The investors are concerned about the presence of fake and dummy accounts on the platform. Your task is to provide data on users, particularly bots, who have liked every single photo on the site, as it is improbable for a normal user to accomplish such a feat.

B.

```
134
135   /*Bots & Fake Accounts:
138
139 •   select user_id ,count(*) as num_likes
140     from likes
141     group by user_id
142     having num_likes = (SELECT COUNT(*) FROM photos);
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

user_id	num_likes
5	257
14	257
21	257
24	257
36	257
41	257

Result 13

There are 13 users who have liked every single photo posted on the site. Using the user id of these bots, they can be removed from the platform to make Instagram a spam free community

Summary:

1. **The top 5 oldest customers who are considered to be the most loyal ones created their accounts in the month of May. Darby was the first one to register**
2. **Only 5 users have never posted a single picture on Instagram as of yet.**
3. **The contest winner is Zack_Kemmer93 whose id is 52 gets the 48 likes on a single photo.**
4. **These are the 5 most commonly used hashtags on the platform**

   i.e. smile, beach, party, fun, concert.
5. **There are 13 bots on the platform that could fill the app with spam and irrelevant content.**
6. **found that Thursdays and Sundays were the days when most users registered.**
7. **On an average, a user was posting 3-4 photos in a year.**

Result:

SQL is a vital skill for individuals in data-driven roles. It enables efficient extraction of information and facilitates the construction and analysis of metrics.

These insights are invaluable for identifying the most loyal customers, pinpointing inactive users, understanding commonly used hashtags, and detecting bots on the site. This learning underscores the significance of considering a multitude of metrics beyond just monetary ones in the business decision-making process.

THANK YOU.