INSTAGRAM USER ANALYTICS

Project report

Shubhangi Chaudhary

Project Description

The project aims to analyze user interactions and engagement with the Instagram app to provide valuable insights that can help a business grow.
The user analysis involves tracking how users engage with a digital product, such as a software application or a mobile app. The insights derived from this analysis can be used by various teams within the business.

Approach

The basic steps involved the creation of the many databases and it was followed by effectively addressing the interaction and relevant behaviour, the focus was on the following:

A) Marketing Analysis:

1. **For the Loyal User Reward-** the five oldest users on Instagram from the database were opted for.
2. **For Inactive User Engagement-** users who have never posted a single photo on Instagram were identified.
3. **For Contest Winner Declaration-** the user with the most likes on a single photo was found and their details were provided to the team.
4. **For Hashtag Research-** the top five most commonly used hashtags on the platform were chosen.
5. **For Ad Campaign Launch-** the day of the week when most users register on Instagram was determined.

B) Investor Metrics:

1. **For User Engagement-** the average number of posts per user on Instagram was determined besided the total number of photos on Instagram divided by the total number of users.
2. **For Bots & Fake Accounts-** users (potential bots) who have liked every single photo on the site, were determined.

Tech-Stack Used

MySQL Workbench and Ms-Excel were used to carry out the following analysis. It allowed for the approach to be simpler and the results could be achieved in a comprehensive and user-friendly way.

Insights

The following insights were gained based on the focused approach:

A) Marketing Analysis:

- **For the Loyal User Reward-** the five oldest users on Instagram from the database were opted for.

```
Loyal User Reward: The marketing team wants to reward the most loyal users, i.e., those who have been using the platform for the longest time.
Task: Identify the five oldest users on Instagram from the provided database.*/
SELECT
    id, username, created_at
FROM
    users
ORDER BY created_at
LIMIT 5;
```

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

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

Inference: The five oldest users were hence determined .

- **For Inactive User Engagement-** users who have never posted a single photo on Instagram were identified.

```
/*Inactive User Engagement: The team wants to encourage inactive users to start posting by sending them promotional emails.
Task: Identify users who have never posted a single photo on Instagram.*/
SELECT
    *
FROM
    users
WHERE
    users.id NOT IN (SELECT
            user_id
        FROM
            photos);
```

id	username	created_at
5	Aniya_Hackett	2016-12-07 01:04:39
7	Kasandra_Homenick	2016-12-12 06:50:08
14	Jadyn81	2017-02-06 23:29:16
21	Rocio33	2017-01-23 11:51:15
24	Maxwell.Halvorson	2017-04-18 02:32:44
25	Tierra.Trantow	2016-10-03 12:49:21
34	Pearl7	2016-07-08 21:42:01
36	Ollie_Ledner37	2016-08-04 15:42:20
41	Mckenna17	2016-07-17 17:25:45
45	David.Osinski47	2017-02-05 21:23:37
49	Morgan.Kassulke	2016-10-30 12:42:31
53	Linnea59	2017-02-07 07:49:34
54	Duane60	2016-12-21 04:43:38
57	Julien_Schmidt	2017-02-02 23:12:48
66	Mike.Auer39	2016-07-01 17:36:15
68	Franco_Keebler64	2016-11-13 20:09:27
71	Nia_Haag	2016-05-14 15:38:50
74	Hulda.Macejkovic	2017-01-25 17:17:28
75	Leslie67	2016-09-21 05:14:01
76	Janelle.Nikolaus81	2016-07-21 09:26:09
80	Darby_Herzog	2016-05-06 00:14:21
81	Esther.Zulauf61	2017-01-14 17:02:34
83	Bartholome.Bernhard	2016-11-06 02:31:23
89	Jessyca_West	2016-09-14 23:47:05
90	Esmeralda.Mraz57	2017-03-03 11:52:27
91	Bethany20	2016-06-03 23:31:53
NULL	NULL	NULL

Inference: Inactive users were hence determined as mentioned above.

- **For Contest Winner Declaration-** the user with the most likes on a single photo was found and their details were provided to the team.

```
/*Contest Winner Declaration: The team has organized a contest where the user with the most likes on a single photo wins.
Task: Determine the winner of the contest and provide their details to the team.*/
SELECT
    users.username,
    photos.id,
    photos.image_url,
    COUNT(*) AS max_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 max_likes DESC
LIMIT 1;
```

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

username	id	image_url	max_likes
Zack_Kemmer93	145	https://jarret.name	48

Inference: The winner was hence located.

- **For Hashtag Research-** the top five most commonly used hashtags on the platform were chosen.

```
/*Hashtag Research: A partner brand wants to know the most popular hashtags to use in their posts to reach the most people.
SELECT
    t.tag_name AS 'name of tag',
    COUNT(p.photo_id) AS hashtag_frequency
FROM
    photo_tags AS p
        INNER JOIN
    tags AS t ON t.id = p.tag_id
GROUP BY t.tag_name
ORDER BY hashtag_frequency DESC
LIMIT 5;
```

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

name of tag	hashtag_frequency
smile	59
beach	42
party	39
fun	38
concert	24

Inference: The 5 most hashtags were found as mentioned above.

- **For Ad Campaign Launch-** the day of the week when most users register on Instagram was determined.

```
/*Ad Campaign Launch: The team wants to know the best day of the week to launch ads.
Task: Determine the day of the week when most users register on Instagram. Provide insights on when to schedule an ad campaign.*/
SELECT
    DAYNAME(created_at) AS day_,
    COUNT(id) AS no_of_registrations
FROM
    users
GROUP BY day_
ORDER BY no_of_registrations DESC
```

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

day_	no_of_registrations
Thursday	16
Sunday	16
Tuesday	14
Saturday	12
Wednesday	13
Monday	14
Friday	15

Inference: Thursdays and Sundays were when most users registered.

B) Investor Metrics:

- **For User Engagement-** the average number of posts per user on Instagram was determined besided the total number of photos on Instagram divided by the total number of users.

```
/*Investor Metrics:
Your Task: Calculate the average number of posts per user on Instagram. Also, provide the total number of photos on Instagram and the total number of users*/
SELECT
    USER_ID, COUNT(ID) AS COUNTS
FROM
    PHOTOS
GROUP BY user_id;-- average post per person

SELECT
    COUNT(*) AS total_ig_images
FROM
    photos;-- total photos
SELECT
    COUNT(username) AS total_ig_users
FROM
    users;-- total users
```

USER_ID	COUNTS
1	5
2	4
3	4
4	3
6	5
8	4
9	4
10	3
11	5
12	4
13	5
15	4
16	4
17	3
18	1
19	2
20	1

total_ig_images
257

total_ig_users
100

Inference: The average post per user and total number of images and users were thus found.

- **For Bots & Fake Accounts-** users (potential bots) who have liked every single photo on the site, were determined.

```
/*Your Task: Identify users (potential bots) who have liked every single photo on the site, as this is not typically possible for a normal user.*/
SELECT
    photo_id,COUNT(user_id) AS Likes
FROM
    likes
GROUP BY photo_id
ORDER BY photo_id
```

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

photo_id	Likes
1	25
2	36
3	38
4	38
5	31
6	31
7	38
8	27
9	31
10	30
11	33
12	29
13	40
14	35
15	34
16	37
17	36

Inference: Bots were possibly absent since no users has liked all images.

Result

The obtained results provided a better insight into the behaviour of the users and helped us better understand and determine different aspects relevant to the needs of the team to optimize their approach and gain satisfactory results.