Instagram User Analytics

Project Description: The aim of the project is to provide a clear insight based on the various data collected from the users, which will help us to improve the user experience and marketing in the Instagram application.

Approach: This project was worked based on the raw data collected from the database, which were used to create a database. This project was executed in SQL, by extracting, filtering and manipulation of the data from the database.

Tech-Stack Used: In this Project, the Tech-Stack we used is MySQL Workbench 8.0 CE. This tool is easy to access and optimized for querying in the database.

Insights:

A) Marketing Analysis:

1. ***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.

   Conclusion: Below provided the list of five oldest users on Instagram from the provided database.

ID	User Name	Creation time
80	Darby_Herzog	06-05-2016 00:14
67	Emilio_Bernier52	06-05-2016 13:04
63	Elenor88	08-05-2016 01:30
95	Nicole71	09-05-2016 17:30
38	Jordyn.Jacobson2	14-05-2016 07:56

   Code:

```
SELECT *
FROM
    users
    order by created_at
    limit 5;
```

Output:

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

2. ***Inactive User Engagement:*** The team wants to encourage inactive users to start posting by sending them promotional emails.

   Conclusion: Below mentioned the list inactive users, who had never posted a single photo on Instagram.

S. No	User id	User name
1	5	Aniya_Hackett
2	7	Kasandra_Homenick
3	14	Jaclyn81
4	21	Rocio33
5	24	Maxwell.Halvorson
6	25	Tierra.Trantow
7	34	Pearl7
8	36	Ollie_Ledner37
9	41	Mckenna17
10	45	David.Osinski47
11	49	Morgan.Kassulke
12	53	Linnea59
13	54	Duane60
14	57	Julien_Schmidt
15	66	Mike.Auer39
16	68	Franco_Keebler64
17	71	Nia_Haag
18	74	Hulda.Macejkovic
19	75	Leslie67
20	76	Janelle.Nikolaus81
21	80	Darby_Herzog
22	81	Esther.Zulauf61
23	83	Bartholome.Bernhard
24	89	Jessyca_West
25	90	Esmeralda.Mraz57
26	91	Bethany20

Code:

```
SELECT *
FROM
        users
        left JOIN
        photos ON  users.id=photos.user_id
        where user_id is null;
```

Output:

id	username	created_at	id	image_url	user_id	created_dat
5	Aniya_Hackett	2016-12-07 01:04:39	NULL	NULL	NULL	NULL
7	Kasandra_Homenick	2016-12-12 06:50:08	NULL	NULL	NULL	NULL
14	Jaclyn81	2017-02-06 23:29:16	NULL	NULL	NULL	NULL
21	Rocio33	2017-01-23 11:51:15	NULL	NULL	NULL	NULL
24	Maxwell.Halvorson	2017-04-18 02:32:44	NULL	NULL	NULL	NULL
25	Tierra.Trantow	2016-10-03 12:49:21	NULL	NULL	NULL	NULL
34	Pearl7	2016-07-08 21:42:01	NULL	NULL	NULL	NULL
36	Ollie_Ledner37	2016-08-04 15:42:20	NULL	NULL	NULL	NULL
41	Mckenna17	2016-07-17 17:25:45	NULL	NULL	NULL	NULL
45	David.Osinski47	2017-02-05 21:23:37	NULL	NULL	NULL	NULL
49	Morgan.Kassulke	2016-10-30 12:42:31	NULL	NULL	NULL	NULL
53	Linnea59	2017-02-07 07:49:34	NULL	NULL	NULL	NULL
54	Duane60	2016-12-21 04:43:38	NULL	NULL	NULL	NULL
57	Julien_Schmidt	2017-02-02 23:12:48	NULL	NULL	NULL	NULL
66	Mike.Auer39	2016-07-01 17:36:15	NULL	NULL	NULL	NULL
68	Franco_Keebler64	2016-11-13 20:09:27	NULL	NULL	NULL	NULL
71	Nia_Haag	2016-05-14 15:38:50	NULL	NULL	NULL	NULL
74	Hulda.Macejkovic	2017-01-25 17:17:28	NULL	NULL	NULL	NULL
75	Leslie67	2016-09-21 05:14:01	NULL	NULL	NULL	NULL
76	Janelle.Nikolaus81	2016-07-21 09:26:09	NULL	NULL	NULL	NULL
80	Darby_Herzog	2016-05-06 00:14:21	NULL	NULL	NULL	NULL
81	Esther.Zulauf61	2017-01-14 17:02:34	NULL	NULL	NULL	NULL
83	Bartholome.Bernhard	2016-11-06 02:31:23	NULL	NULL	NULL	NULL
89	Jessyca_West	2016-09-14 23:47:05	NULL	NULL	NULL	NULL
90	Esmeralda.Mraz57	2017-03-03 11:52:27	NULL	NULL	NULL	NULL
91	Bethany20	2016-06-03 23:31:53	NULL	NULL	NULL	NULL

3. ***Contest Winner Declaration:*** The team has organized a contest where the user with the most likes on a single photo win.

Conclusion: Below provided the details of the winner of the contest

User id	User Name	Photo id	Number of Likes	image_url
52	Zack_Kemmer93	145	48	https://jarret.name

Code:

```
with Noof_likes as (
select photo_id, count(photo_id) as Num_likes
from likes
    group by photo_id
    order by Num_likes desc)
select photo_id, Num_likes, image_url, user_id, username
from Noof_likes
    inner join photos
    on Noof_likes.photo_id=photos.id
    inner join users
    on photos.user_id=users.id
    Group by photo_id
    Order by Num_likes desc
    limit 1;
```

Output:

user_id	username	photo_id	Num_likes	image_url
52	Zack_Kemmer93	145	48	https://jarret.name

4. ***Hashtag Research:*** A partner brand wants to know the most popular hashtags to use in their posts to reach the most people.

   Conclusion: Below provided the list of five most commonly used hashtags on the platform.

Hashtag name	No. of times Hashtag used
smile	59
beach	42
party	39
fun	38
concert	24

Code:

```
with hashtags as (
    select * from tags
    inner join photo_tags
    on tags.id=photo_tags.tag_id
```

)

```
Select tag_name, count(tag_name) as Noof_tags
from hashtags
        Group by tag_name
        Order by Noof_tags desc
        limit 5;
```

Output:

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

5. ***Ad Campaign Launch:*** The team wants to know the best day of the week to launch ads.

Conclusion: The best day of the week to launch ads is Thursday.

Code:

```
with date_info as        (
        select id, username, created_at, dayname(created_at) as days
        from users
        )
select days, count(days) as numof_ids
from date_info
        group by days
        order by numof_ids desc
        limit 1;
```

Output:

days	numof_ids
Thursday	16

B) Investor Metrics:

1. ***User Engagement:*** Investors want to know if users are still active and posting on Instagram or if they are making fewer posts.

   Conclusion: The average number of posts per user on Instagram is 2.57.

   Code:

```
SELECT
        COUNT(image_url) AS totalphotos,
        COUNT(DISTINCT username) AS totalusers,
        COUNT(image_url) / COUNT(DISTINCT username) AS avg_photosperuser
FROM
        photos
        RIGHT JOIN
        users ON photos.user_id = users.id;
```

   Output:

totalphotos	totalusers	avg_photosperuser
257	100	2.5700

2. ***Bots & Fake Accounts:*** Investors want to know if the platform is crowded with fake and dummy accounts.

   Conclusion: Below mentioned the list of potential bots on the site

S. No	User id	No of photos liked	User name
1	5	257	Aniya_Hackett
2	14	257	Jaclyn81
3	21	257	Rocio33
4	24	257	Maxwell.Halvorson
5	36	257	Ollie_Ledner37
6	41	257	Mckenna17
7	54	257	Duane60
8	57	257	Julien_Schmidt
9	66	257	Mike.Auer39
10	71	257	Nia_Haag
11	75	257	Leslie67
12	76	257	Janelle.Nikolaus81
13	91	257	Bethany20

Code:

```
SELECT
        user_id, COUNT(photo_id) AS no_oflikes, username
FROM
        likes
        INNER JOIN
        users ON likes.user_id = users.id
        GROUP BY user_id
        HAVING no_oflikes = 257;
```

Output:

user_id	no_oflikes	username
5	257	Aniya_Hackett
14	257	Jaclyn81
21	257	Rocio33
24	257	Maxwell.Halvorson
36	257	Ollie_Ledner37
41	257	Mckenna17
54	257	Duane60
57	257	Julien_Schmidt
66	257	Mike.Auer39
71	257	Nia_Haag
75	257	Leslie67
76	257	Janelle.Nikolaus81
91	257	Bethany20

Result:

1. Successfully collected data, data cleaning and analysed the data to provide valuable insights.
2. Increased my problem solving and analytical skills.
3. Provided Hands on experience in SQL using real world raw data.
4. Based on analysis, the marketing team need to create more promotional emails to encourage usage of the Instagram, as the quarter of the registered users are not using the Instagram.
5. The insights derived from these data are valuable, as they can support us to make data driven operational and marketing decisions.