Instagram User Analytics

Description:

- Project Environment:
- As a data analyst working with the product team at Instagram.
- My role involves analysing user interactions and engagement with the Instagram app to provide valuable insights that can help the business grow.
- Here, the Project Analysis involves:
  - a] Tracking how users engage with a digital product, such as a software application or a mobile app.
  - b] The insights derived from this analysis can be used by various teams within the business. For example,
    a. The marketing team might use these insights to launch a new campaign,
    b. The product team might use them to decide on new features to build, and
    c. The development team might use them to improve the overall user experience.
- The Final insights are then used by teams across the business to launch a new marketing campaign, decide or discover new updated features to build / add into an App. Also, to track the success of the app by measuring user engagement and improve the experience altogether, by helping the business to grow.

Approach:

- In this project, I am using SQL and MySQL Workbench as tool to analyse Instagram user data and answer questions posed by the management team.
- Use SQL queries to extract the required information from the database, ensuring the accuracy and efficiency of the queries.
- The Final insights will help the product manager and the rest of the team to make informed decisions about the future direction of the Instagram app.
- Approach used to complete this Project:
  a. here, we create a "Database" & "table's" as per dataset given to us, for this Project [Instagram User Analytics].
  b. Also, Utilise SQL to analyse data from created tables & find Result / Solution's for the given problem statements / tasks

- Project Solution: [writing a query to find solution for given problem statement.]
  [A]. Marketing Analysis:
  1) Identify the five oldest users on Instagram from the provided database.
     [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."]

```sql
select *
from users
order by created_at limit 5;
```

2) identify users who have never posted single photo on Instagram.
[Inactive User Engagement: The team wants to encourage inactive users to start posting by sending them promotional emails.]

```sql
select *
from users u
left join photos p
on u.id = p.user_id
where p.image_url is null
order by u.id;
```

3) determine the winner of the contest and provide their details to the team.
[Contest Winner Declaration: The team has organized a contest where the user with the most likes on a single photo win.]

```sql
-- first we calculate total likes count from database:

select users.id, users.username, count(likes.user_id) as Total_likes
from likes
inner join photos on likes.photo_id = photos.id
inner join users on photos.user_id = users.id
group by likes.photo_id, users.username, users.id
order by Total_likes desc ;

-- now, we find winner with respect to total likes counts:

select users.id, users.username, count(likes.user_id) as Total_likes
from likes
inner join photos on likes.photo_id = photos.id
inner join users on photos.user_id = users.id
group by likes.photo_id, users.username, users.id
order by Total_likes desc limit 3;
```

4) identify and suggest top five most commonly used hashtags to use in their posts to reach the most people.
[Hashtag Research: A partner brand wants to know the most popular hashtags to use in their posts to reach the most people.]

```sql
select t.tag_name , count( pt.photo_id ) as hashtags
from photo_tags pt
join tags t on pt.tag_id = t.id
group by t.tag_name
order by hashtags desc;
```

5) determine the day of the week when most users register on Instagram. provide insights on when to schedule an ad campaign.
[Ad Campaign Launch: The team wants to know the best day of the week to launch ads.]

```sql
select date_format((created_at), '%W') as days, count(username) as total_user_register
from users
group by days
order by total_user_register desc;
```

[B]. Investor Metrics:

1. Calculate the average number of posts per user on Instagram. Also, provide the total number of photos on Instagram divided by the total number of users.
[User Engagement: Investors want to know if users are still active and posting on Instagram or if they are making fewer posts.]

```sql
-- first we calculate post per user, on instagram:

select u.id as user_ID, count(p.id) as post_per_user
from users u
left join photos p on u.id = p.user_id
group by u.id;


-- Now, to calculate average number of post per user,
--                         we first calculate, "total number of photos on Instagram", & "the total number of users".
--                                 [ select * from users, photos; ]


with base as(
            select u.id as user_ID, count(p.id) as Totalpost_per_user
            from users u
            left join photos p on u.id = p.user_id
            group by u.id
            )
select sum(Totalpost_per_user) as total_photos, count(user_ID) as total_users, round((sum(Totalpost_per_user) / count(user_ID)), 1) as Photo_per_user
from base;
```

2. Identify users (potential bots) who have liked every single photo on the site, as this is not typically possible for a normal user.
   [Bots & Fake Accounts: Investors want to know if the platform is crowded with fake and dummy accounts.]

```sql
-- calculate likes per user on instagam:

select u.username, count(l.photo_id) as likes_per_user
from users u
left join likes l on u.id = l.user_id
group by u.username
order by likes_per_user;


-- Now, Identify (potential bots) users, who have liked every single photo on the site. [which normally not the case.]

with base as(
    select u.username, count(l.photo_id) as likes_per_user
    from likes l
    left join users u on u.id = l.user_id
    group by u.username
    order by likes_per_user
    )
select username, likes_per_user
from base                              -- Using where clause: to compare likes_per_person with total like_photo count.
where likes_per_user = (select(max(likes.photo_id)) from likes)
order by username;
```

Tech-Stack Used:

Administration – Server Status:

MySQL [Sever] Workbench – local community version.

version – 8.0.41 (MySQL Community Server GPL)

Insights:

- After analysing database, our insightful findings help market team: To find loyal User & inactive users on Instagram to use separate strategies to promote and enlarge userbase.
- Also, the insights on most used and popular hashtags, help in ad-campaigns and brand promotions to promote & reach the most people.
- Insights help us to find out “Bots & Fake Accounts”, if the Instagram platform is crowded with fake and dummy accounts.

Insights & Result-Solution for given problem Statements:

[A]. Marketing Analysis:

1) Identify the five oldest users on Instagram from the provided database.
   [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."]

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

2) identify users who have never posted single photo on Instagram.
   [Inactive User Engagement: The team wants to encourage inactive users to start posting by sending them promotional emails.]

id	username	created_at	id	image_url	user_id	created_dat
5	Aniya Hackett	2016-12-07 01:01:39	NULL	NULL	NULL	NULL
7	Kassandra Homenick	2016-12-12 06:50:08	NULL	NULL	NULL	NULL
14	Jadyn81	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 23:12:18	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) determine the winner of the contest and provide their details to the team.
   [Contest Winner Declaration: The team has organized a contest where the user with the most likes on a single photo win.]

id	username	Total_likes
52	Zack_Kemmer93	48
46	Malinda_Streich	43
65	Adelle96	43

4) identify and suggest top five most commonly used hashtags to use in their posts to reach the most people.
[Hashtag Research: A partner brand wants to know the most popular hashtags to use in their posts to reach the most people.]

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

5) determine the day of the week when most users register on Instagram. provide insights on when to schedule an ad campaign.
[Ad Campaign Launch: The team wants to know the best day of the week to launch ads.]

days	total_user_register
Thursday	16
Sunday	16
Friday	15
Tuesday	14
Monday	14
Wednesday	13
Saturday	12

[B]. Investor Metrics:

1.1) Calculate the average number of posts per user on Instagram.
[User Engagement: Investors want to know if users are still active and posting on Instagram or if they are making fewer posts.]

user_ID	username	post_per_user
1	Kenton_Kirlin	5
2	Andre_Purdy85	4
3	Harley_Lind18	4
4	Arely_Bogan63	3
5	Aniya_Hackett	0
6	Travon.Waters	5
7	Kasandra_Homenick	0
8	Tabitha_Schamberger11	4
9	Gus93	4
10	Presley_McClure	3
11	Justina.Gaylord27	5
12	Dereck65	4
13	Alexandro35	5
14	Jaclyn81	0
15	Billy52	4

user_ID	username	post_per_user
16	Annalise.McKenzie16	4
17	Norbert_Carroll35	3
18	Odessa2	1
19	Hailee26	2
20	Delpha.Kihn	1
21	Rocio33	0
22	Kenneth64	1
23	Eveline95	12
24	Maxwell.Halvorson	0
25	Tierra.Trantow	0
26	Josianne.Friesen	5
27	Darwin29	1
28	Dario77	4
29	Jaime53	8
30	Kaley9	2

user_ID	username	post_per_user
31	Aiyana_Hoeger	1
32	Irwin.Larson	4
33	Yvette.Gottlieb91	5
34	Pearl7	0
35	Lennie_Hartmann40	2
36	Ollie_Ledner37	0
37	Yazmin_Mills95	1
38	Jordyn.Jacobson2	2
39	Kelsi26	1
40	Rafael.Hickle2	1
41	Mckenna17	0
42	Maya.Farrell	3
43	Janet.Armstrong	5
44	Seth46	4
45	David.Osinski47	0

user_ID	username	post_per_user
46	Malinda_Streich	4
47	Harrison.Beatty50	5
48	Granville_Kutch	1
49	Morgan.Kassulke	0
50	Gerard79	3
51	Mariano_Koch3	5
52	Zack_Kemmer93	5
53	Linnea59	0
54	Duane60	0
55	Meggie_Doyle	1
56	Peter.Stehr0	1
57	Julien_Schmidt	0
58	Aurelie71	8
59	Cesar93	10
60	Sam52	2

user_ID	username	post_per_user
61	Jayson65	1
62	Ressie_Stanton46	2
63	Elenor88	4
64	Florence99	5
65	Adelle96	5
66	Mike.Auer39	0
67	Emilio_Bernier52	3
68	Franco_Keebler64	0
69	Karley_Bosco	1
70	Erick5	1
71	Nia_Haag	0
72	Kathryn80	5
73	Jaylan.Lakin	1
74	Hulda.Macejkovic	0
75	Leslie67	0

user_ID	username	post_per_user
76	Janelle.Nikolaus81	0
77	Donald.Fritsch	5
78	Colten.Harris76	5
79	Katarina.Dibbert	1
80	Darby_Herzog	0
81	Esther.Zulauf61	0
82	Aracely.Johnston98	2
83	Bartholome.Bernhard	0
84	Alysa22	2
85	Milford_Gleichner42	2
86	Delfina_VonRueden68	9
87	Rick29	4
88	Clint27	11
89	Jessyca_West	0
90	Esmeralda.Mraz57	0

user_ID	username	post_per_user
87	Rick29	4
88	Clint27	11
89	Jessyca_West	0
90	Esmeralda.Mraz57	0
91	Bethany20	0
92	Frederik_Rice	3
93	Willie_Leuschke	2
94	Damon35	1
95	Nicole71	2
96	Keenan.Schamberger60	3
97	Tomas.Beatty93	2
98	Imani_Nicolas17	1
99	Alek_Watsica	3
100	Javonte83	2

1.2) provide the total number of photos on Instagram divided by the total number of users.

total_photos	total_users	Photo_per_user
256	100	2.6

2) Identify users (potential bots) who have liked every single photo on the site, as this is not typically possible for a normal user. [**Bots & Fake Accounts**: Investors want to know if the platform is crowded with fake and dummy accounts.]

username	likes_per_user
Aniya_Hackett	257
Bethany20	257
Duane60	257
Jaclyn81	257
Janelle.Nikolaus81	257
Julien_Schmidt	257
Leslie67	257
Maxwell.Halvorson	257
Mckenna17	257
Mike.Auer39	257
Nia_Haag	257
Ollie_Ledner37	257
Rocio33	257

Results:

Results we achieve:

- After analysing database, our insightful findings help market team: To find loyal User & inactive users on Instagram to use separate strategies to promote and enlarge userbase.
- Also, the insights on most used and popular hashtags, help in ad-campaigns and brand promotions to promote & reach the most people.
- Insights help us to find out “Bots & Fake Accounts”, if the Instagram platform is crowded with fake and dummy accounts.

Drive Link:

This Project help me to achieve expertise in MySQL data analytics.

Also, give me hand-on Expertise & Experience in industry standard project-dataset handling, with different business problem statements/ requirements.