Instagram User Analytics

Project Description:

- THIS PROJECT IS TO TRACK HOW USERS ENGAGE & INTERACT IN THE PLATFORM.
- IDENTIFYING THE MOST LOYAL USERS ON THE PLATFORM.
- IDENTIFYING INACTIVE USERS.
- IDENTIFYING MOST LIKED PHOTO POSTED BY A USER.
- IDENTIFYING TOP 5 HASHTAGS USED ON THE PLATFORM.
- LAUNCHING AD CAMPAIGN
- IDENTIFYING AVERAGE USER POST ON INSTAGRAM, TOTAL NO OF PHOTOS & TOTAL NO OF USERS.
- IDENTIFYING BOT & FAKE USERS.

Approach:

- We have to analyze the data provided in the dataset.
- Identify the tables on which we require to perform operations.
- Utilize the data in the given tables and perform operations to get desired results.

Tech-Stack Used

- Software used: My SQL version 8.0 by Oracle.
- Reason for using:
  - Freeware
  - Hassle free installation
  - UI is easy to use & user friendly

Insights:

- As I am new to SQL I learn a lot of new SQL commands.
- Formulation of SQL queries as per requirement.

Results:

A.Marketing:

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

- Query: `select id,username,created_at from users order by created_at limit 5;`
- Output:

	id int	username varchar	created_at timestamp
1	80	Darby_Herzog	2016-05-06 00:14:21
2	67	Emilio_Bernier52	2016-05-06 13:04:30
3	63	Elenor88	2016-05-08 01:30:41
4	95	Nicole71	2016-05-09 17:30:22
5	38	Jordyn.Jacobson2	2016-05-14 07:56:26

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

- Query: select username from users left join photos on users.id=photos.user_id where photos.id is null;
- Output:

	username varchar
1	Aniya_Hackett
2	Kasandra_Homenick
3	Jaclyn81
4	Rocio33
5	Maxwell.Halvorson
6	Tierra.Trantow
7	Pearl7
8	Ollie_Ledner37
9	Mckenna17
10	David.Osinski47
11	Morgan.Kassulke
12	Linnea59
13	Duane60
14	Julien_Schmidt
15	Mike.Auer39
16	Franco_Keebler64
17	Nia_Haag
18	Hulda.Macejkovic
19	Leslie67
20	Janelle.Nikolaus81
21	Darby_Herzog
22	Esther.Zulauf61
23	Bartholome.Bernhard
24	Jessyca_West
25	Esmeralda.Mraz57
26	Bethany20

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

- Query: select users.username,photos.id,photos.image_url,count(*) as total_likes from likes

  join photos on photos.id=likes.photo_id

  join users on users.id=likes.photo_id

  group by photos.id

  order by total_likes desc limit 1;

- Output:

	username varchar	id int	image_url varchar	total_likes bigint
1	Kaley9	30	http://kenny.com	41

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

- Query: select tags.tag_name,photo_tags.tag_id,count(tag_id)
  count from photo_tags , tags
  where tags.id = photo_tags.tag_id
  group by tags.tag_name,photo_tags.tag_id order by count desc limit 5;
- Output:

	tag_name varchar	tag_id int	count bigint
1	smile	21	59
2	beach	20	42
3	party	17	39
4	fun	13	38
5	concert	18	24

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

- Query: select weekday(created_at), count(created_at) from users group by weekday(created_at);
- Output:

	weekday(created_at) int	count(created_at) bigint
1	3	16
2	6	16
3	1	14
4	5	12
5	2	13
6	0	14
7	4	15

- **Note:** 0 = Monday, 1 = Tuesday, 2 = Wednesday, 3 = Thursday, 4 = Friday, 5 = Saturday, 6 = Sunday.
- On Thursday & Sunday most users registered

B. Investor Metrics:

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/total number of users

- Query:select count(image_url) , count(distinct(user_id)), count(image_url)/count(distinct(user_id)) as Average from photos;
- Output:

	count(image_url) bigint	count(distinct(user_id)) bigint	Average_post newdecimal
1	257	74	3.4730

- Total no of Photos: 257
- Total no of users: 74
- Average Post: 3.473

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).

- Query: `select L.user_id, U.username, count(L.photo_id) X from likes L inner join users U on L.user_id = U.id group by L.user_id, U.username order by X desc;`
- Output:

	user_id int	username varchar	Photos_liked bigint
1	75	Leslie67	257
2	21	Rocio33	257
3	24	Maxwell.Halvorson	257
4	91	Bethany20	257
5	36	Ollie_Ledner37	257
6	41	Mckenna17	257

	user_id int	username varchar	Photos_liked bigint
7	14	Jaclyn81	257
8	76	Janelle.Nikolaus81	257
9	54	Duane60	257
10	57	Julien_Schmidt	257
11	66	Mike.Auer39	257
12	5	Aniya_Hackett	257
13	71	Nia_Haag	257

- Bot Users are the users who have liked all 257 images