



# *Project Description*

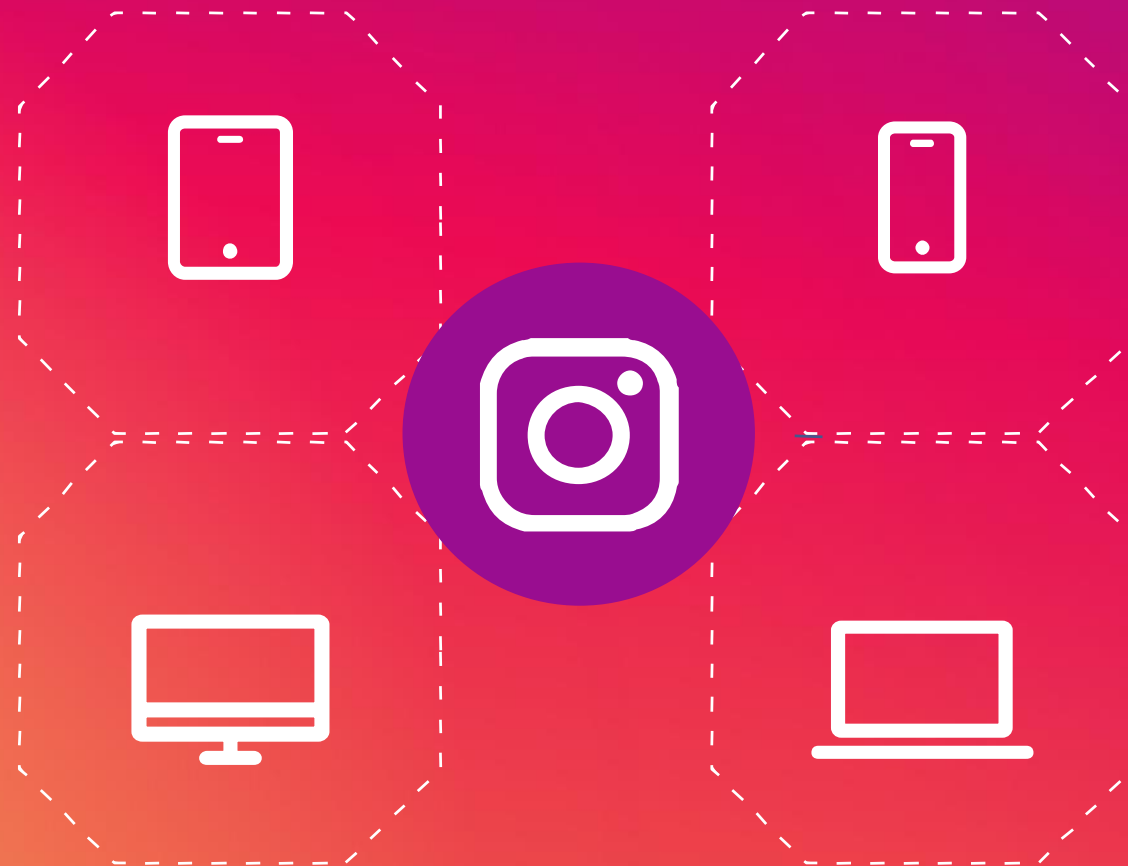
- for this project Instagram provided us with a dataset containing user analytics. Using this dataset, we must generate some insights and answer some of the questions asked by Instagram's management team.
- The company will use these insights to launch new marketing campaigns, choose which features to include in apps, evaluate the success of the apps by looking at user engagement, and generally improve the user experience while assisting in business expansion.

# Approach

- First of all I learned which tasks I needed to perform and which kind of insights the management team was expecting, from the description that is given.
- I used MySQL Workbench to import and inspected the dataset that was provided.
- Then I started to run the required queries, and if there were any errors in the code, I modified it until it was completely error-free.
- Finally, I ensured that the produced insights met the management team's expectations and answered all of their inquiries.

# 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.  
**Your Task:** Find the 5 oldest users of Instagram from the database provided

Query  
SELECT \*  
FROM users  
ORDER BY created\_at  
LIMIT 5;

**Insight:**

Below is the list of the top 5 most loyal users that can be rewarded for using the platform for the longest period of time. Darby Herzog should be rewarded the most because he/she has been using the account for the longest

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. 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 ROW_NUMBER() OVER (ORDER BY users.id) AS sr_no,username
FROM photos
RIGHT JOIN users ON users.id = photos.user_id
WHERE photos.id IS NULL;
```

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

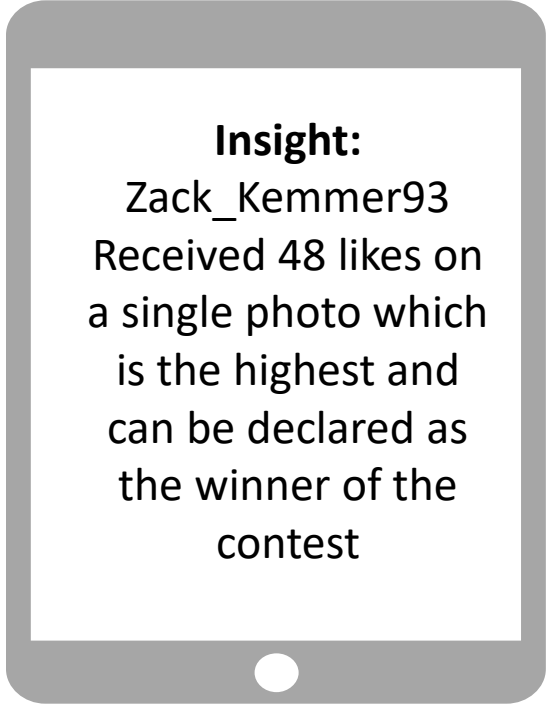
**Insight:**

Here is the list of all the inactive users who have never posted a single photo on Instagram. The sales team should send them a promotional email and remind them to post their 1st photo.

**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 = photos.user_id  
GROUP BY photos.id  
ORDER BY total_likes  
DESC LIMIT 1;
```



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

**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 tag_name,  
COUNT(tag_name) AS total  
FROM tags  
JOIN photo_tags  
ON tags.id = photo_tags.tag_id  
GROUP BY tags.id  
ORDER BY total  
DESC LIMIT 5;
```

**Insight:**

Following is the list of the top 5 most commonly used hashtags. #smile can be used to reach more people as it is the most-used hashtag among all of them.

tag_name	total
smile	59
beach	42
party	39
fun	38
concert	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 day_of_the_week,  
count(day_of_the_week)as created_day  
from(SELECT *,weekday(created_at)as day_of_the_week from users)a  
group by day_of_the_week  
order by created_day desc;
```

**Insight:**

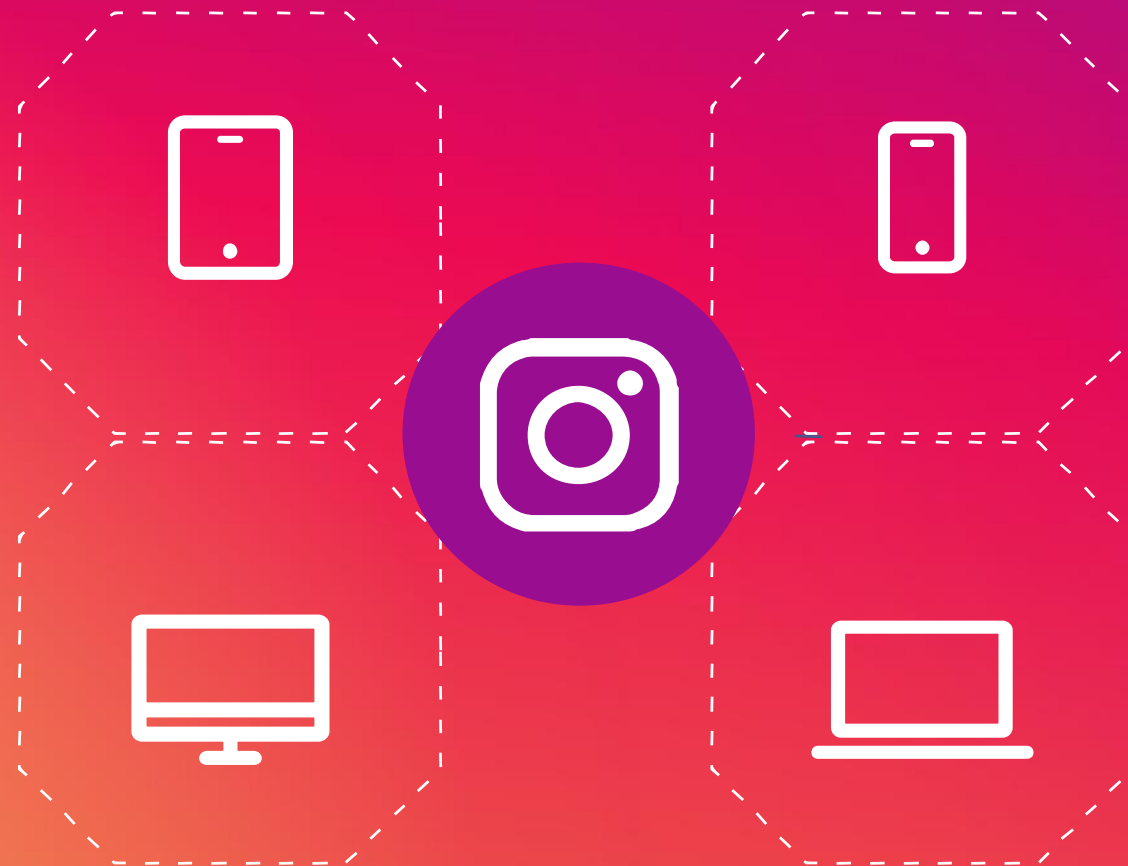
The team can launch the ad campaign on the 3<sup>rd</sup>, 6<sup>th</sup> and 4<sup>th</sup> day of the week as the most number of users registered on these days



day_of _week	created_day
3	16
6	16
4	15
1	14
0	14
2	13
5	12

## B) Investor Metrics:

Our investors want to know if Instagram is performing well and is not becoming redundant like Facebook, they want to assess the app on the following grounds



**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: Total number of photos on Instagram**

```
SELECT  
ROUND((SELECT COUNT(*) FROM PHOTOS));
```

Total Photos
257

**Query: Total number of users**

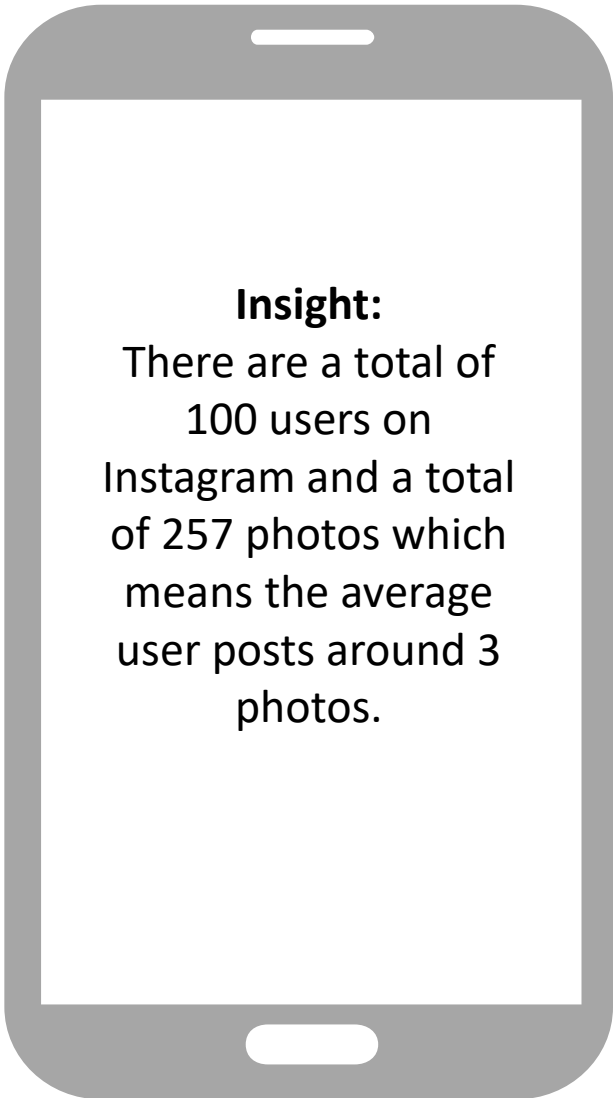
```
SELECT  
ROUND((SELECT COUNT(*)FROM USERS));
```

Total Users
100

**Query: For average user posts on Instagram**

```
SELECT  
ROUND(SELECT COUNT(*)FROM PHOTOS)/(SELECT COUNT(*)FROM USERS);
```

average user posts on Instagram
2.57



**Insight:**

There are a total of  
100 users on  
Instagram and a total  
of 257 photos which  
means the average  
user posts around 3  
photos.

**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 username,  
COUNT(\*) AS num\_likes  
FROM users  
INNER JOIN likes ON users. id = likes.user\_id  
GROUP BY likes. user\_id  
HAVING num\_likes = (SELECT COUNT(\*) FROM photos);

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

**Insight:**

This is the list of users who have liked every single photo on Instagram which is not possible for a normal user to do so. Hence we can consider them as bots

# Learnings

- Through this project I learned how to handle practical questions within an organization.
- I observed how data science can help us to solve even the most difficult questions in the organization.
- This project also helped me to enhance my knowledge of SQL and challenged my problem-solving skills.
- Overall by completing this project I came to know how actual data analysts analyze user data and derive insights out of it.

## Tech-Stack Used

- All of the analysis part is done using the My SQL Workbench Version 8.0.32
- All the presentation-related work is done using Microsoft PowerPoint Professional Plus 2021

## Drive link

[https://drive.google.com/drive/folders/16OgLFkxNdgMt3UhIjObfZo-yXlMtWKVf?usp=share\\_link](https://drive.google.com/drive/folders/16OgLFkxNdgMt3UhIjObfZo-yXlMtWKVf?usp=share_link)