

# **INSTAGRAM USER ANALYTICS**

# PROJECT DESCRIPTION

- This Project is Regarding the data analytics of Instagram user's where we extract the useful insights from the data that we get from the management.
- We need to provide the required and useful insights from question's given the management and investors.

# MY APPROACH

- Firstly I have checked the data completely and by reviewing the data that we got from the management.
- Understanding each question got from the management and investors.
- Visualizing the ways so that I get the required metrics from that data.
- Which function's are needed to be used for extracting the specific insight from data.



# INSIGHTS

- There are various insights I gained by doing the analysis of the data.
- Where I used various SQL functions In an efficient way.
- I came to know about metrics like the average per user's, how to identify the bots in platform.
- Which day the user's are most active ,most used hashtag's , most likes on single photo of a particular user.
- Where we have to get data from various databases and join them extract important metrics.



# TECH-STACK USED

- To analyze the data and to gain insights from the data I have used the My SQL Workbench

# MARKETING

## MOST LOYAL USERS

USERNAME	CREATED_AT
Darby_Herzog	06-05-2016 00:14
EMILIO_BERNIER52	06-05-2016 13:04
ELENOR88	08-05-2016 01:30
NICOLE71	09-05-2016 17:30
JORDYN.JACOBSON2	14-05-2016 07:56

- The above user's are most loyal users according to the data because these user's are created there accounts ahead of time before many user's present in the data.
- The above user's are oldest user's of the platform we can that see by ascending order of account created date.
- The marketing team can reward those user's as most loyal user's of the platform.



## SQL Queries Used :-

```
SELECT username, created_at  
From Users  
ORDER by created_at ASC  
LIMIT 5;
```

## INACTIVE USERS :-

ID	USERNAME
5	Aniya_Hackett
7	Kasandra_Homenick
14	Jaclyn81
21	Rocio33
24	Maxwell.Halvorson
25	Tierra.Trantow
34	Pearl7
36	Ollie_Ledner37
41	Mckenna17

- The above user's are inactive user's in data provided where those didn't post any photo till date.
- We know that metric by going through the particular user's photo\_id and image url.
- If we get the image url and user's photo\_id then he is an active user (or) we get image url and photo\_id as NULL then he is an inactive user.
- So the team can send them the promotional email's to post there 1<sup>st</sup> photo for above user's.





## SQL Queries Used :-

```
SELECT username, created_at  
From Users  
ORDER by created_at ASC  
LIMIT 5;
```

## CONTEST WINNER:-

ID	USERNAME	PHOTO ID	HIGHEST LIKES
52	Zack_Kemmer93	145	48

- The contest Winner is Zack\_kemmer93 who got 48 likes on the single photo.

## SQL Queries Used :-

```
SELECT  
users.id,username,likes.photo_id,COUNT(*)  
highest_likes  
FROM ig_clone.users users  
JOIN ig_clone.photos photos  
ON users.id = photos.user_id  
JOIN ig_clone.likes likes  
ON photos.id= likes.photo_id  
GROUP by 3  
ORDER by 4 DESC  
LIMIT 1
```

## TOP HASHTAG'S USED :-

TAG NAME	MOST USED
smile	59
beach	42
party	39
fun	38
concert	24

- Smile,beach,party,fun and concert are most used hastag's on photo's by platform user's.
- If we use the same hashtag's in brand post's the reach will be higher and most people will interact with the post.

## SQL Queries Used :-

```
SELECT tag_name,COUNT(*)  
most_used  
FROM ig_clone.tags  
JOIN ig_clone.photo_tags  
ON id=tag_id  
GROUP BY tag_name  
ORDER BY most_used desc  
LIMIT 5
```

## AD LAUNCH CAMPAIGN:-

REGISTERED_DAY	COUNT(*)
Thursday	16
Sunday	16

- Most user's are registered on Thursday and Sunday on the platform.
- So when we have to launch an AD campaign we can do that on that day.

## SQL Queries Used :-

```
SELECT dayname(created_at) regsiteer_day,  
count(*)  
FROM ig_clone.users  
GROUP BY 1  
ORDER BY 2 DESC  
LIMIT 2
```

# INVESTOR METRICS

## USER ENGAGEMENT

AVG_USER_POST
2.57

- There are 100 users on platform and those user's posted 257 posts on the platform.
- The average posts on platform by users would be posts/users so that will be 2.57.



## SQL Queries Used :-

```
SELECT  
  (SELECT COUNT(*) FROM ig_clone.photos)/(SELECT  
  COUNT(*) FROM ig_clone.users)  
AS avg_user_post
```

# BOTS & FAKE ACCOUNTS:-

BOTS	LIKES ON PHOTOS
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

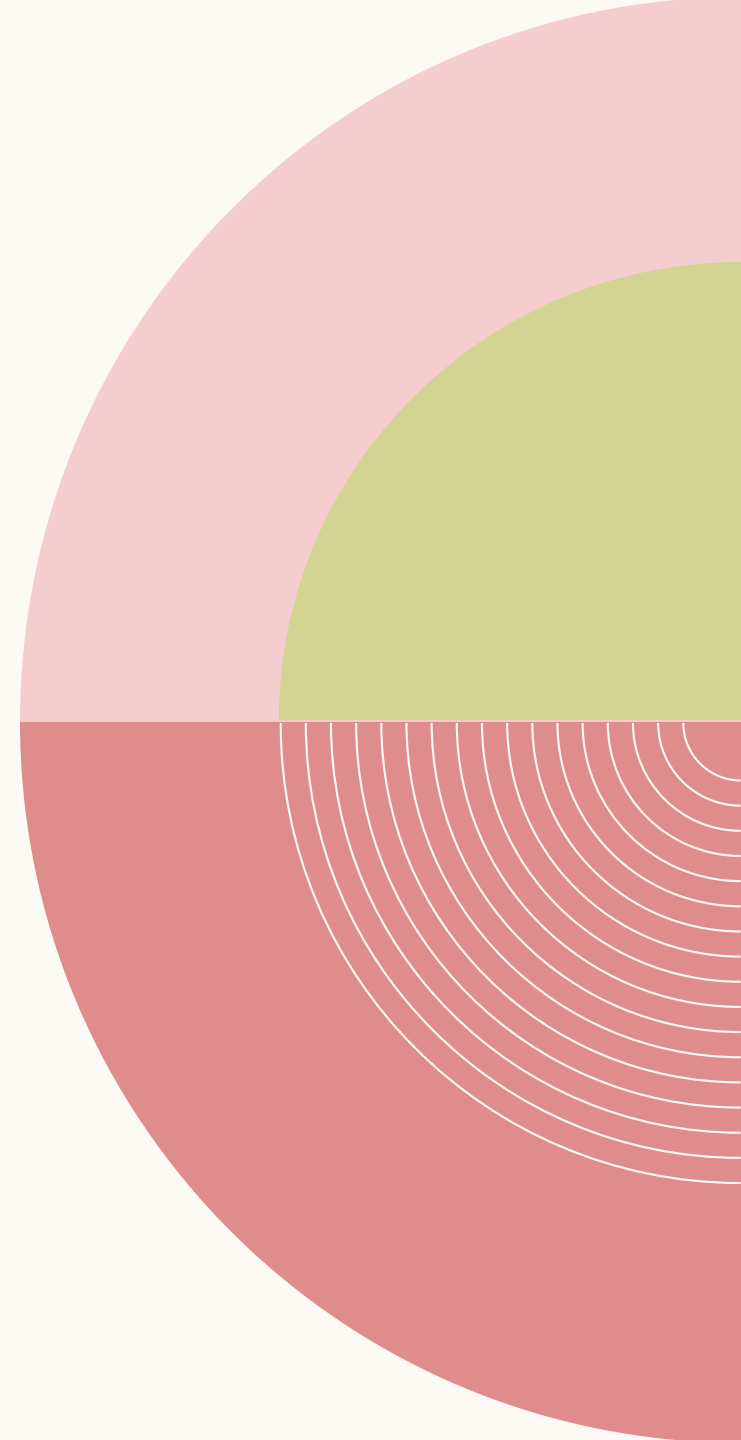
- The above user's are fake (or) bots as per given condition these user's likes the every post on the platform posted by the user's.

## SQL Queries Used :-

```
SELECT users.username bots,  
COUNT(likes.photo_id) likes_on_photos  
FROM ig_clone.likes likes  
JOIN ig_clone.users users  
ON likes.user_id = users.id  
GROUP BY 1  
HAVING likes_on_photos = MAX(likes.photo_id);
```

# RESULT

- This project is very helpful for me to apply the SQL language to perform some logical operations to perform on the given data.
- Where I analyzed the data and I gained very useful insights from that data about the user's and there activity on the platform.
- Used join functions , aggregate function and sorting functions on the data to gain the useful hidden metrics in data.
- I learned lot of practical applications of SQL and it's functions and that enhanced my SQL and Analytic Skills.



The background features a large, light cream-colored circle on the left and a large, light pink circle on the right. These two circles overlap in the center. The area where they overlap is filled with a series of thin, white, concentric circular lines that radiate from the center of the overlap. The top and bottom edges of the image are framed by a solid dark blue color.

**THANK YOU**