# INSTAGRAM USER ANALYTICS

### **AGENDA**



### CREATING DATABASE

- We can create the database by using the following command.
- SYNTAX : CREATE DATABASE DATABASE\_NAME;

- In order to use that particular database we need to use the command
- SYNTAX : USE DATABASE\_NAME;
- In order to see databases we have to use the command
- SYNTAX : SHOW DATABASES;

```
create database ig_clone;
use ig_clone;
show databases;
```

#### A) Marketing Analysis:

**QUERY 1**: Identify the five oldest users on Instagram from the provided database.

#### **PROJECT DESCRIPTION: LOYAL USER REWARD**

The marketing team wants to encourage the loyal customers and also reward them to retain them. So, in order to identify those customers we need to identify the top most old users of Instagram.

- **SELECT**: We choose the columns id, username, and created\_at.
- FROM: Get the data from the "users" table.
- **ORDER BY**: Sort the results by created\_at column in ascending order.
- **LIMIT**: Limit to show only the first 5 rows.

#### •**QUERY 1:**

```
id, username, created_at
FROM
users
ORDER BY created_at
LIMIT 5;
```

	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. Jacobson 2	2016-05-14 07:56:26
	NULL	NULL	NULL

#### A) Marketing Analysis:

**QUERY 2:** Identify users who have never posted a single photo on Instagram.

#### **PROJECT DESCRIPTION: INACTIVE USER ENGAGEMENT**

The team wants to encourage the inactive user to start posting there photos so inorder to encourage them they want to send the promotional emails to the inactive users.

- •**SELECT:** We choose the columns users.id and users.username.
- •FROM: We get the data from the users table.
- •LEFT JOIN: To combine the users table with the photos table on the condition that users.id matches photos.user\_id.
- •WHERE: To filter the results to include only those rows where photos.user\_id is NULL.

#### •QUERY 2:

```
SELECT

users.id, users.username

FROM

users

LEFT JOIN

photos ON users.id = photos.user_id

WHERE

photos.user_id IS NULL;
```

190		
id	username	
7	Kasandra_Homenick	
14	Jaclyn81	
21	Rocio33	
24	Maxwell.Halvorson	
25	Tierra.Trantow	
34	Pearl7	
36	Ollie_Ledner37	
41	Mckenna 17	
45	David.Osinski47	
49	Morgan.Kassulke	
53	Linnea59	
54	Duane60	
57	Julien_Schmidt	
66	Mike. Auer 39	
68	Franco Keebler64	
71	Nia Haag	
74	Hulda.Macejkovic	
75	Leslie67	
76	Janelle Nikolaus 81	
80	Darby_Herzog	
81	Esther.Zulauf61	
83	Bartholome.Bernhard	
89	Jessyca_West	
90	Esmeralda.Mraz57	
91	Bethany20	

#### A) Marketing Analysis:

**QUERY 3:** Determine the winner of the contest and provide their details to the team.

#### **PROJECT DESCRIPTION: CONTEST WINNER DECLARATION**

The team wants to organize a contest where the user whose has more number of likes for a single post is declared as winner.

- •SELECT: Choose columns users.username, users.id, users.created\_at
- •FROM: users table.
- •**JOIN:** photos on users.id = photos.user\_id
- •WHERE: photos.id matches the most liked photo, determined by a subquery on the likes table.

#### •QUERY 3:

```
SELECT
    users.username, users.id, users.created_at
FROM
    users
        JOIN
    photos ON users.id = photos.user_id
WHERE
    photos.id = (SELECT
            photo_id
        FROM
            likes
        GROUP BY photo_id
        ORDER BY COUNT(photo_id) DESC
       LIMIT 1);
```

	username	id	created_at
١	Zack_Kemmer93	52	2017-01-01 05:58:22

#### A) Marketing Analysis:

**QUERY 4**: Identify and suggest the top five most commonly used hashtags on the platform.

#### **PROJECT DESCRIPTION: HASHTAG RESEARCH**

The brand wants to know the most popular hashtag that are used ,to keep for there posts to reach more people.

- •SELECT:Choose column tags.tag\_name
- •FROM: tags table
- •JOIN: photo\_tags on tags.id = photo\_tags.tag\_id
- •GROUP BY: tags.id
- •ORDER BY: Count of photo\_tags.tag\_id in descending order
- •LIMIT: Show only the top 5 results

#### •**QUERY 4:**

```
Tags.tag_name

FROM

tags

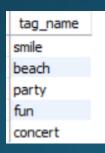
JOIN

photo_tags ON tags.id = photo_tags.tag_id

GROUP BY tags.id

ORDER BY COUNT(photo_tags.tag_id) DESC

LIMIT 5;
```



#### A) Marketing Analysis:

**QUERY 5:** To determine the day of the week when most users register on Instagram. Provide insights on when to schedule an ad campaign.

#### **PROJECT DESCRIPTION:** AD CAMPAIGN LAUNCH

The brand wants to know the on which days more number of people are active on platform so that they can launch the ads that reach many audience.

- •SELECT: DAYNAME(created\_at) as dayname and COUNT(\*) as regst
- •FROM: users table.
- •GROUP BY: dayname
- •ORDER BY: regst in descending order
- •LIMIT: Show only the top result

#### •**QUERY 5**:

```
DAYNAME(created_at) AS dayname, COUNT(*) AS regst
FROM
users
GROUP BY dayname
ORDER BY regst DESC
LIMIT 1;
```

#### **RESULT:**

dayname regst Thursday 16

#### **B)** Investor Metrics:

**QUERY 1**: To 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.

#### **PROJECT DESCRIPTION: USER ENGAGEMENT**

Investors want to know whether the user are still active by posting the photos or posting fewer.

- SELECT: column:AVG(nop) as avg\_posts\_per\_user
- •FROM: A subquery that:
- •SELECT: user\_id and COUNT(\*) as nop
- •FROM: photos
- •OUP BY: user\_id

#### **B)** Investor Metrics:

**QUERY 1:** To 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.

#### **PROJECT DESCRIPTION: USER ENGAGEMENT**

Investors want to know whether the user are still active by posting the photos or posting fewer.

#### **APPROACH:**

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

- •SELECT: The result of dividing the total number of photos by the total number of users as average\_posts\_per\_user
- •FROM: Two subqueries that:
- •SELECT: COUNT(\*) from photos
- •SELECT: COUNT(\*) from users

#### •QUERY 1:

```
SELECT

AVG(nop) AS avg_posts_per_user

FROM

(SELECT

user_id, COUNT(*) AS nop

FROM

photos

GROUP BY user_id) AS user_post_counts;
```

```
SELECT

((SELECT

COUNT(*)

FROM

photos) / (SELECT

COUNT(*)

FROM

users)) AS average_posts_per_user;
```

#### **RESULT:**

avg\_posts\_per\_user
3.4730

average\_posts\_per\_user 2.5700

#### **B)** Investor Metrics:

**QUERY 2:** To Identify users who have liked every single photo on the site, as this is not typically possible for a normal user.

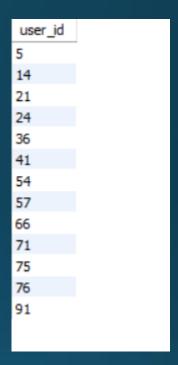
#### PROJECT DESCRIPTION: BOTS AND FAKEACCOUNTS

Investors want to know whether the platform consists of fake accounts or bots.

- SELECT: Choose columns user\_id
- •FROM: likes table.
- •GROUP BY: grouped by user\_id column
- •HAVING: The count of likes by each user matches the total count of photos, determined by a subquery on the photos table.

#### •**QUERY 2:**

#### 



### TECH-STACK USED

**SOFTWARE:** MYSQL WORKBENCH 8.0 CE

#### **IMPORTANCE OF MYSQL:**

- 1. Security: MySQL provides security features like user authentication and data encryption to protect data that database hold.
- 2. Speed: It executes the query very fast and high performance, even with large datasets.
- 3. Simplicity: It is easy to learn and use.
- 4. Accuracy: It ensures data integrity with features like constraints and transaction management.
- 5. Accessibility: We can accessible across platforms and integrates with various applications and programming languages.

### **INSIGHTS**

- We can easily filter the people based on different criteria, which makes our easy.
- We can identify the top most loyal users and most users active on which day for promotion .
- To declare the result for the contest.
- Identifying the patterns help to encourage more people to involve based on there interest.
- o To identify the most used Hashtags and the most liked posts.

### CONCLUSION

- o Data analytics helps to identify the user behavior, patterns
- O Which allows the marketing team to target the right audience for different campaigns.
- o It guides the product team to develop which kind of feature.
- o It guides the development team to look after overall performance and experience.
- So, It is very important to analysis the data and predict the things for better prosper for the company.

## THANK YOU